*Article*

# The Difficulties in Emotion Regulation Scale–Parent Report: A Psychometric Investigation Examining Adolescents With and Without ADHD

Assessment
1–20

DOI: 10.1177/1073191118792307
journals.sagepub.com/home/asm
SAGE

**Nóra Bunford[1,2,3], Anne E. Dawson[3], Steven W. Evans[3], A. Raisa Ray[3], Joshua M. Langberg[4], Julie Sarno Owens[3], George J. DuPaul[5], and Darcey M. Allan[3]**

## Abstract

Emotion dysregulation is associated with attention deficit/hyperactivity disorder (ADHD) and confers risk for behavior problems and functional impairment; however, there is little guidance on best practices for measurement in adolescents. We developed a parent-report version of the Difficulties in Emotion Regulation Scale (DERS-P)). Evidence of reliability and validity was evaluated in a large community online sample (Study 1: $n = 978$; $M_{age} = 13.52$ years; $SD = 1.93$) and in two samples of adolescents with ADHD (Study 2, Sample 1: $n = 78$; $M_{age} = 12.12$ years, $SD = 0.91$; Sample 2: $n = 206$; $M_{age} = 15.35$ years; $SD = 0.85$). A four-factor solution of the DERS-P was obtained in Study 1 and confirmed in Study 2, with factors demonstrating acceptable internal consistency. The community sample was generally rated as less dysregulated than the ADHD samples. Support was obtained for convergent, concurrent, and incremental validity evidence. These findings provide preliminary evidence for the DERS-P as a psychometrically sound parent-report measure of emotion dysregulation in 11- to 17-year-old adolescents.

## Keywords

adolescence, emotion regulation/dysregulation, measurement, rating scale, parent report, ADHD

Despite variability across definitions, emotion regulation can be conceptualized as an individual's ability to modulate (1) the speed with which and degree to which the physiological, experiential, and behavioral expression of an emotion escalates; (2) the intensity of the physiological, experiential, and behavioral expression of an emotion; and (3) the speed with which and degree to which physiological, experiential, and behavioral expression of an emotion deescalates in a manner congruent with an optimal level of functioning (Bunford, Evans, & Wymbs, 2015, p. 188). Emotion *dys*regulation (ED) is an individual's inability to exercise any or all aspects of the emotion regulation modulatory processes, to such a degree that the inability results in the individual failing to successfully adapt to and meet environmental challenges and goals and thus functioning meaningfully below his or her baseline (Bunford, Evans, & Wymbs, 2015, p. 188).

ED is a transdiagnostic characteristic associated with externalizing (e.g., oppositional defiant disorder), internalizing (e.g., anxiety, depression), and, as defined here, neurodevelopmental disorders (Bunford, Evans, & Wymbs, 2015; Mazefsky et al., 2013) such as attention deficit/hyperactivity disorder (ADHD; Bunford, Evans, Becker, & Langberg, 2015; Bunford, Evans, et al., 2017; Bunford, Evans, & Langberg, 2014; Graziano & Garcia, 2016; Okado & Mueller, 2016; Sjöwall, Backman, & Thorell, 2015). Although whether ED is a core feature or merely an associated feature of ADHD remains debated (Barkley, 2010), there is a growing body of work indicating that ED is associated with negative outcomes such as alcohol problems (Bunford, Wymbs, Dawson, & Shorey, 2017) and smoking (Mitchell et al., 2012), and with functional impairment in the academic and social domains (Bunford, Evans, Becker, et al., 2015; Maedgen & Carlson, 2000; McQuade, Penzel, Silk, & Lee, 2016; Melnick & Hinshaw, 2000)

[1]Institute of Cognitive Neuroscience and Psychology, Budapest, Hungary
[2]Eötvös Loránd University, Budapest, Hungary
[3]Ohio University, Athens, OH, USA
[4]Virginia Commonwealth University, Richmond, VA, USA
[5]Lehigh University, Bethlehem, PA, USA

**Corresponding Author:**
Nóra Bunford, Institute of Cognitive Neuroscience and Psychology, Hungarian Academy of Sciences, Research Centre for Natural Sciences, Magyar tudósok körútja 2, Budapest 1117, Hungary.
Email: nb243610@ohio.edu

among children, adolescents, and adults with ADHD. These findings underscore the importance of ED in association with ADHD across the life span; however, there is little guidance on best practices for measuring ED in this population, especially in adolescents (Bunford, Evans, et al., 2017; Bunford, Evans, & Wymbs, 2015).

ED in children with ADHD has been evaluated using multiple methodologies, ranging from the most micro to the most macro level (Morris & Cuthbert, 2012). Examples include genetic (Merwood et al., 2014) physiological (Bunford, Evans, et al., 2017; Musser et al., 2011; Musser, Galloway-Long, Frick, & Nigg, 2013), observational (C. Lee et al., 2017; Maedgen & Carlson, 2000; Melnick & Hinshaw, 2000), parent report and self-report on rating scale (Seymour, Chronis-Tuscano, Iwamoto, Kurdziel, & MacPherson, 2014) and ecological momentary assessment (Rosen & Epstein, 2010; Rosen, Epstein, & Van Orden, 2013) methods.

Conversely, considerably less is known about ED in adolescents with ADHD and what is known is based exclusively on studies wherein ED was assessed via self-report on rating scales (Bunford et al., 2014; Bunford, Evans, Becker, et al., 2015; Seymour et al., 2012). These limitations are problematic, for the following reasons. First, upward extensions of findings obtained with children to information about adolescents is unfounded, in part because the development of emotion processing and regulation capacities during adolescence has unique aspects and challenges. Relative to children, adolescents experience more intense and labile emotions (Silk, Steinberg, & Morris, 2003), during a period when asynchronicity between brain development, behavioral, cognitive, and emotional self-regulatory skill development, and environmental demands for self-regulation is at its peak (Steinberg, 2005). Deficits or delays in acquisition of emotion processing and regulation skills can thus be especially debilitating for adolescents, with this debilitation likely exacerbated in the context of psychiatric disorders or symptoms. For example, in the context of ADHD, above and beyond symptoms, a number of manifestations of ED have been shown to predict comorbidities and functional impairments (Bunford et al., 2014; Bunford, Evans, Becker, et al., 2015; Seymour et al., 2014). In addition, in youth with ADHD, there are some differences and similarities in the aspects of ED that are relevant to children relative to adolescents. For example, emotional lability, negativity, and reactivity are less prevalent among younger than in older youth (Graziano & Garcia, 2016) but across children and adolescents, emotional inflexibility and slow return to baseline have been shown to predict indices of observed and parent- and self-reported social impairment (Bunford et al., 2014; Melnick & Hinshaw, 2000).

Second, evident from their definitions, emotion regulation and dysregulation are multidimensional phenomena, both in terms of the corresponding experiential or phenomenological experience and of the observable or outward manifestation. Yet the noted exclusive reliance in adolescents on self-report precludes comprehensive measurement of ED. Despite the importance of using comparable assessment methods across two or more respondents for gaining contextual information (i.e., contexts in which a characteristic does or does not occur; De Los Reyes et al., 2015) and such methods being considered best-practice in evaluating mental health in youth (Youngstrom, Choukas-Bradley, Calhoun, & Jensen-Doss, 2015), the absence of valid measures of ED across informants precludes parallel assessment across adolescents and their parents (Mash & Hunsley, 2005). Furthermore, because of increases in cognitive maturity and the internal nature of emotions, self-report may be an appropriate method for measuring adolescent ED (Rohrbeck, Azar, & Wagner, 1991; Soto, John, Gosling, & Potter, 2008; Walden, Harris, & Catron, 2003). However, youth with ADHD exhibit a substantial response bias and tendency to use a dichotomous and positively skewed response style when reporting their emotional experiences and regulation (Rosen et al., 2013), challenging the view that self-report is sufficient, especially for adolescents with the disorder (see, e.g., Maedgen & Carlson, 2000; Melnick & Hinshaw, 2000, for evidence of observable behavioral correlates of ED in youth with ADHD). Rather, to complement data obtained via self-report measures on the experiential experience of ED, development and evaluation of parallel parent-report measures that capture observable manifestations of ED is needed.

One thoroughly evaluated and widely used self-report measure of ED is the Difficulties in Emotion Regulation Scale (DERS; Gratz & Roemer, 2004). Although the DERS was developed and its psychometric properties were initially evaluated with adult samples (Gratz & Roemer, 2004), evidence for acceptable psychometric properties of the measure has been established in a variety of adolescent samples, including good internal consistency, test–retest reliability, adequate construct, and predictive validity (Adrian et al., 2009; Bunford et al., 2014; Bunford, Evans, Becker, et al., 2015; Vasilev, Crowell, Beauchaine, Mead, & Gatzke-Kopp, 2009; Weinberg & Klonsky, 2009), with robust correlations with psychological problems reflecting ED (Weinberg & Klonsky, 2009) and with physiological measures of ED (Vasilev et al., 2009). There are additional advantages of the DERS, including that the measure is based on an integrative conceptualization of emotion regulation, one that also encompasses the definition adopted in the current study. Specifically, that conceptualization involves not only the modulation of emotional arousal and its various physiological, experiential, and behavioral expressions but also the acceptance, awareness, and understanding of emotions, as well as ability to act in desired ways even in the face of strong emotions (Gratz & Roemer, 2004). Findings also support the transdiagnostic utility of the measure, insofar as they indicate that the DERS or subscales predict disorder

status (e.g., generalized anxiety disorder: Salters-Pedneault, Roemer, Tull, Rucker, & Mennin, 2006; posttraumatic stress disorder: McDermott, Tull, Gratz, Daughters, & Lejuez, 2009; Tull, Barrett, McMillan, & Roemer, 2007), past childhood maltreatment (Soenke, Hahn, Tull, & Gratz, 2010), nonsuicidal self-injury (Perez, Venta, Garnaat, & Sharp, 2012), and depression (Ehring, Fischer, Schnülle, Bösterling, & Tuschen-Caffier, 2008).

Despite these advantages to the DERS, there is some concern regarding the stability of its factor structure, potentially due to problems with one of its factors that consists of reverse-coded items (Bardeen, Fergus, Hannan, & Orcutt, 2016; Bardeen, Fergus, & Orcutt, 2012; D. J. Lee, Witte, Bardeen, Davis, & Weathers, 2016). In the first study on the factor structure of the DERS (Gratz & Roemer, 2004), although a six- and a seven-factor solution was also identified, the authors retained the six-factor solution, as it was more interpretable: Awareness (six items; indicative of awareness or acceptance of one's emotions); Clarity (five items; indicative of clarity or knowledge about one's emotions); Goals (five items; indicative of difficulty with goal-directed behavior when upset); Impulse (six items; indicative of difficulty with behavioral control when experiencing negative emotions); Nonacceptance (six items; indicative of negative secondary emotions); and Strategies (eight items; indicative of a belief that no strategy can decrease negative emotion). Similarly, Weinberg and Klonsky (2009) found evidence for both a six- and a seven-factor solution and chose to retain the six-factor solution, arguing that the seventh factor may have represented an artifact of item format rather than content. In an attempt to determine a source of these problems related to the factor structure of the DERS, Bardeen and colleagues made note of and tested problems with the Awareness subscale that comprises only reverse-coded items (Bardeen et al., 2012; Bardeen et al., 2016; D. J. Lee et al., 2016). Bardeen et al. (2016) forward coded all reverse-coded items, submitted them to exploratory factor analysis (EFA) in one sample, and then to confirmatory factor analysis (CFA) in another sample and identified five factors, evidence of potential problems with the Awareness subscale and the reverse-coded items.

Despite these inconsistencies across these studies, there is a body of research that, although not unequivocally and not with regard to all aspects of psychometric properties, does support the six-factor structure of the DERS, including use of the Awareness subscale and reverse-coded items (Coutinho, Ribeiro, Ferreirinha, & Dias, 2010; Hervás, 2011). This line of research includes studies conducted with large samples across cultures and languages (Mitsopoulou, Kafetsios, Karademas, Papastefanakis, & Simos, 2013 [$n = 780$]; Neumann, van Lier, Gratz, & Koot, 2010 [$n = 870$]) as well as across racial groups ($n =$ Caucasian 446, African American 424, and Asian American 180) and sexes ($n =$ male 256 and female 794; Ritschel, Tone, Schoemann, & Lim, 2015). Beyond considerations regarding factor structure, the Awareness subscale is meaningfully, albeit sometimes less strongly (e.g., Neumann et al., 2010) correlated with the other DERS subscales (Fowler et al., 2014; Giromini, Velotti, De Campora, Bonalume, & Cesare Zavattini, 2012; Perez et al., 2012; Ritschel et al., 2015) and is related to relevant outcomes, including delinquency (Neumann et al., 2010), differentiating cocaine-dependent patients and community-controlled volunteers (Fox, Axelrod, Paliwal, Sleeper, & Sinha, 2007) and differentiating individuals with alcohol dependence during abstinence from social drinkers (Fox, Hong, & Sinha, 2008).

Taken together, although the results of some studies do bring into question the established factor structure of the self-report DERS and raise questions about the Awareness subscale, the results of other studies provide evidence supporting the validity of that factor structure. Extended examination of the psychometric properties of this self-report measure of ED in adolescents, including to cross-informant consistency and convergent and incremental validity would bolster confidence in these findings. Simultaneously, development and evaluation of parallel measures would fulfill the need to establish reliable and valid tools to assess ED that do not rely exclusively on self-report and thus increase our ability to gather multi-informant data on an important transdiagnostic construct.

## Current Study

To begin addressing the gap in the literature, we developed a parent-report form of the DERS (DERS-P; Gratz & Roemer, 2004)[1]. Despite potential problems with the self-report version, initiating a line of research examining a parent-report form is warranted in light of the considerable clinical promise and utility of such parent-rated version. Our aim in the current research was to evaluate the psychometric properties—including evaluating evidence for reliability and validity—of the DERS-P among adolescents with and without ADHD.

Our first two aims were to examine (1) the factor structure and (2) the internal consistency of the DERS-P. These aims are addressed in Study 1 using a large online national sample of parents of adolescents. Subsequently, we wanted to examine the evidence of (3) the convergent validity of the DERS-P as indexed by its association with two self-report measures of ED, (4) the concurrent validity of the DERS-P as indexed by its association with theoretically relevant constructs such as anxiety and depression, and (5) the incremental validity of the DERS-P as indexed by the degree to which it predicts, beyond the self-report DERS, outcomes associated with ED. These aims were addressed in Study 2 using a community sample and a sample of parents of adolescents with ADHD and their children.

We predicted that (1) the self-report and parent-report forms will not have the same factor structure given that certain aspects of ED are directly observable, some can only be inferred, and others are largely internal. This hypothesis is consistent with differences in the psychometric properties of parent- and child-report versions of rating scales of other characteristics. For example, others have reported a different factor structure between parent- and child-report versions of measures of callous unemotional traits (Gao & Zhang, 2016) and of health quality of life (Steele, 2008). As the current research was aimed at evaluating the psychometric properties of the DERS-P, our design was exploratory and thus we did not formulate hypotheses about the other aspects of such properties— internal consistency or convergent, concurrent, and incremental validity.

## Study 1: Exploration of Factor Structure

### Method

*Participants.* Participants were 978 Amazon's Mechanical Turk (i.e., MTurk; www.mturk.com) workers who reported to be U.S. residents and parents of youth aged 11 to 17 years. There were participants from each of the 50 states and the District of Columbia. Participants were predominantly married (80%) and mothers (57.8%). Although most participants were White (77.3%), other racial groups were represented (8.4% Black, 5.2% Latinx, and 5.0% Asian). Annual total family income was assessed in ranges (in US$), and were generally well distributed with most parents reporting income considered generally within the middle class (23.1% made $25,001-45,000, 21.0% made $45,001-65,000, 21.6% made $65,001-85,000, 15.3% made $85,001-105,000, 8.2% made $105, 001-125,000, 4.9% made $125,001-145,000, with the remainder 5.9% making >$145,001 annually). Concerning the mother (or female guardian) of the child, 1.1% did not complete high school, 13.3% had high school diploma, 18.8% obtained some college but no degree, 17.8% received additional technical training of 2-year degree, 33.5% obtained a 4-year degree, and 13.7% obtained higher than a 4-year degree. Concerning the father (or male guardian) of the child, 2.1% did not complete high school, 14.6% had high school diploma, 14.8% obtained some college but no degree, 16.6% received additional technical training of 2-year degree, 33.8% obtained a 4-year degree, and 12.2% obtained higher than a 4-year degree. Thus, a greater percentage of parents reported having obtained higher educational degree than the general population (29.8% of adults >25 years of age report having a 4-year degree or higher; U.S. Census Bureau, 2016).

Youths' ages ranged from 11 to 17 years ($M = 13.52$; $SD = 1.93$). There was a slightly higher percentage of boys (51.2%), and like parents, most adolescents were White (72.4%), with other racial groups represented (8.0% Black, 5.4% Latinx, 4.3% Asian, and 8.0% multiracial). A small percentage (12.1%) of parents indicated that their child has special needs eligible for school-based services. Of these youth, 55.1% had specials needs due to a neurodevelopmental disorder (e.g., autism spectrum disorder, ADHD), 10.2% due to academic problems (e.g., learning disorder, speech delay), 22.05% due to medical problems (e.g., cystic fibrosis, diabetes), and 29.7% due to emotional and behavioral disorders (e.g., anxiety, depression), not exclusively.

*Procedures.* Participants were recruited through MTurk, the largest online crowdsourcing platform used frequently and effectively for (clinical) research (Chandler & Shapiro, 2016). On MTurk, individuals (i.e., referred to as workers) can complete tasks for compensation (Sheehan & Pittman, 2016). Although MTurk samples are not as representative of the general population as national probability samples, they are more representative of the U.S. population than typical convenience samples (e.g., college or community samples) and workers produce reliable and valid data (Chandler & Shapiro, 2016; Sheehan & Pittman, 2016).

MTurk qualifiers were used to limit eligible respondents to individuals who were U.S. residents and had worker approval ratings of ⩾85% (an MTurk index of the quality of participants' responding). Study measures were combined into a single survey, which was advertised as an academic survey about emotion regulation. To qualify for survey completion, respondents first had to provide digital acknowledgment of informed consent. Then, they underwent a brief eligibility screener, wherein they were asked whether they had siblings, were married, and were a parent (and, if so, the ages of their children). Questions about siblings and marriage were used as distractor questions so that respondents would not be aware of qualifying answers. In sum, 3,715 individuals consented to participate; of these, 90.2% reported that they had siblings, 58.7% reported they were married, and 65.3% reported that they were parents. Of parents, 43.8% reported having children between the ages 11 and 17 years, thus passing initial eligibility. Of these, 979 completed the survey. On examination of responses, one respondent was deleted for reporting that he or she was not a resident of the United States. Data collection was completed within approximately 16 hours after posting the survey to MTurk. All measures were parent-report scales. Participants who completed the survey were compensated $2.00 through MTurk payments. This research was approved by the Ohio University Social–Behavioral Institutional Review Board.

*Measures.* The DERS-P was developed by rewording items of the original self-report DERS, which, as noted, has demonstrated evidence of reliability and validity in adult and adolescent samples (Gratz & Roemer, 2004; Weinberg &

Klonsky, 2009). The measure was adapted to assess parental perceptions of adolescent ED without a specified time frame, similar to the self-report measure. For example, the item on the self-report DERS, "I am clear about my feelings" was changed on the DERS-P to "My child is clear about his/her feelings." Items are rated on a 5-point scale (1 = *almost never*, 5 = *almost always*), with higher scores indicating greater ED. The factor structure of the DERS-P will be evaluated in this study. In the current sample, the internal consistency across all 36 DERS-P items was acceptable ($\alpha = .84$).

*Data Analytic Plan.* To address Aim 1, we examined the number and content of the latent factors that represent items included in the DERS-P. Procedures consistent with evidence-based guidelines for scale development (DeVellis, 2016) and identical to those employed with the self-report DERS (Gratz & Roemer, 2004) were followed. Accordingly, we conducted an EFA using all 36 items, using principal axis factoring with promax oblique rotation with Kaiser normalization. We excluded items with poor or dual loadings ($<.40$ on any subscale or $>.40$ on two or more subscales) and further relied on criteria indicating that items with a loading of less than .32 as well as with cross-loadings less than .15 difference from an item's highest factor loading should be eliminated (Brady, Evans, Berlin, Bunford, & Kern, 2012; Worthington & Whittaker, 2006).

To determine factor retention, we examined eigenvalues above 1 and the elbow in the Scree Plot (Costello & Osborne, 2005). To confirm the EFA results based on the Kaiser–Guttman criterion, parallel analysis and the minimum average partial test were conducted. After deletion of an item, the EFA was conducted again until no items were indicated for deletion. Additionally, to address Aim 2, each subscale was assessed for evidence of internal consistency (considered acceptable if $\alpha \geqslant .70$; Nunnally, 1978).

### Results and Discussion

Results of the first EFA indicated seven items (i.e., 4, 5, 9, 17, 20, 31, 34; see Table 1 for item descriptions) to be deleted given inadequate loading on any extracted factor (i.e., $<.40$) through the EFA iterations. Deleting these items also made conceptual sense, as they predominantly pertain to adolescents' beliefs or thoughts (e.g., *believes, is confused, has no idea, has difficulty making sense out of*), which are difficult for any observer, including a parent, to assess, and would thus contribute to greater rater-specific bias (i.e., rater-specific error within each respective item unrelated to the latent factor).

In the final EFA, suitability for factor analysis was examined (Kaiser–Meyer–Olkin [KMO] = 0.959; Bartlett's test = 17,734.29, $p < .001$), and indicated the data were appropriate for factoring. Examining the eigenvalues, four factors had an eigenvalue $>1.0$ and accounted for 57.49% of the variance. On examination of the elbow, the scree plot also appeared to indicate a four-factor solution, and all items adequately loaded onto a factor (see Table 1). Results of the parallel analysis suggested a three-factor solution and those of the minimum average partial test suggested a four-factor solution. It should be noted that although the parallel analysis suggested a three-factor solution, the results fell just below the cut-off for suggesting a four-factor solution (i.e., for the four-factor solution, the eigenvalue for the sample was 1.2347, whereas the eigenvalue for the simulative sample was 1.2387). Given that both the Kaiser–Guttman criterion and the minimum average partial tests suggested a four-factor solution and the parallel analysis was just below the cut-off for suggesting a four-factor solution, a decision was made to select the four-factor solution.

In comparison with the adolescent self-report 36-item scale, two fewer factors emerged in the parent-report version. The original Subscales 3 (Impulse) and 5 (Strategies, with exception of Item 30) include items related to the ability to maintain control over behavioral expression of emotions even in face of strong emotions or to implement strategies to reduce negative emotions—now loaded onto the DERS-P Factor 1. We labeled this factor *Catastrophize*, as many items are related to losing control over and feeling overwhelmed by negative emotions. Most of the items on the self-report DERS that loaded onto Subscale 1 (Nonacceptance) loaded onto the DERS-P Factor 2 (with the addition of Item 30). The Nonacceptance subscale is indicative of a tendency to exhibit negative secondary emotions (Gratz & Roemer, 2004). Accordingly, we labeled this factor *Negative Secondary*. All items that remained from the self-report DERS Subscales 4 (Awareness) and 6 (Clarity), representative of awareness or clarity of emotions, loaded onto the DERS-P Factor 3. We labeled this factor *Attuned* (with greater scores reflecting less awareness or clarity of emotions). Finally, all remaining items from the self-report DERS Subscale 2 (Goals) loaded onto the DERS-P Factor 4. These items are related to difficulty with engaging in goal-directed behavior (e.g., concentrating, focusing, and completing tasks) in the face of strong emotions. As such, we labeled this factor *Distracted*. All factors demonstrated acceptable internal consistency: Factor 1 (Catastrophize; 12 items; $\alpha = .93$), Factor 2 (Negative Secondary; 7 items; $\alpha = .88$), Factor 3 (Attuned; 6-items; $\alpha = .88$), and Factor 4 (Distracted; 4 items; $\alpha = .90$).

## Study 2: Validation of Factor Structure

### Method

*Participants.* In addition to the national online sample of parents described above, this study also included two samples of adolescent participants with ADHD ($n$ = 284

**Table 1.** DERS-P Items and Factor Loadings Following Exploratory Factor Analysis.

| | | Factor Loadings | | | |
|---|---|---|---|---|---|
| OF | Items (DERS-P) | 1 | 2 | 3 | 4 |
| 6 | 1. My child is clear about his/her feelings | .034 | .065 | **.696** | −.023 |
| 4 | 2. My child pays attention to how he/she feels | −.065 | −.02 | **.836** | .04 |
| 3 | 3. My child experiences his/her emotions as overwhelming and out of control | **.764** | .036 | −.087 | −.089 |
| 6 | 4. My child has no idea how he/she is feeling | — | — | — | — |
| 6 | 5. My child has difficulty making sense out of his/her feelings | — | — | — | — |
| 4 | 6. My child is attentive to his/her feelings | −.001 | −.019 | **.792** | −.009 |
| 6 | 7. My child knows exactly how he/she is feeling | −.010 | .007 | **.759** | .116 |
| 4 | 8. My child cares about what he/she is feeling | −.008 | .020 | **.711** | −.074 |
| 6 | 9. My child is confused about how he/she feels | — | — | — | — |
| 4 | 10. When my child is upset, he/she acknowledges his/her emotions | −.005 | .016 | **.690** | −.055 |
| 1 | 11. When my child is upset, he/she becomes angry with him/herself for feeling that way | .289 | **.533** | −.040 | −.077 |
| 1 | 12. When my child is upset, he/she becomes embarrassed for feeling that way | .062 | **.732** | −.041 | −.066 |
| 2 | 13. When my child is upset, he/she has difficulty getting work done | | | | **.780** |
| 3 | 14. When my child is upset, he/she becomes out of control | **.954** | −.076 | −.003 | −.130 |
| 5 | 15. When my child is upset, he/she believes that he/she will remain that way for a long time | **.536** | .167 | −.061 | .122 |
| 5 | 16. When my child is upset, he/she believes that he/she will end up feeling very depressed | **.466** | .321 | −.030 | .002 |
| 4 | 17. When my child is upset, he/she believes that his/her feelings are valid and important | — | — | — | — |
| 2 | 18. When my child is upset, he/she has difficulty focusing on other things | −.018 | −.073 | −.002 | **.910** |
| 3 | 19. When my child is upset, he/she feels out of control | **.821** | .026 | −.032 | .035 |
| 2 | 20. When my child is upset, he/she can still get things done | — | — | — | — |
| 1 | 21. When my child is upset, he/she feels ashamed with him/herself for feeling that way | −.073 | **.831** | .016 | −.037 |
| 5 | 22. When my child is upset, he/she knows that he/she can find a way to eventually feel better | **.497** | −.049 | .318 | −.041 |
| 1 | 23. When my child is upset, he/she feels like he/she is weak | .109 | **.560** | .070 | .039 |
| 3 | 24. When my child is upset, he/she feels like he/she can remain in control of his/her behaviors | **.503** | −.090 | .257 | .029 |
| 1 | 25. When my child is upset, he/she feels guilty for feeling that way | −.181 | **.911** | .031 | −.051 |
| 2 | 26. When my child is upset, he/she has difficulty concentrating | −.112 | .028 | −.008 | **.946** |
| 3 | 27. When my child is upset, he/she has difficulty controlling his/her behaviors | **.796** | −.070 | .027 | .056 |
| 5 | 28. When my child is upset, he/she believes that there is nothing he/she can do to make him/herself feel better | **.567** | .188 | .029 | .052 |
| 1 | 29. When my child is upset, he/she becomes irritated with him/herself for feeling that way | .022 | **.634** | .009 | .114 |
| 5 | 30. When my child is upset, he/she starts to feel very bad about him/herself | .132 | **.617** | .026 | .085 |
| 5 | 31. When my child is upset, he/she believes that wallowing in it is all he/she can do | — | — | — | — |
| 3 | 32. When my child is upset, he/she loses control over his/her behaviors | **.857** | −.065 | −.024 | −.017 |
| 2 | 33. When my child is upset, he/she has difficulty thinking about anything else | .183 | .001 | −.010 | **.681** |
| 4 | 34. When my child is upset, he/she takes time to figure out what he/she is really feeling | — | — | — | — |
| 5 | 35. When my child is upset, it takes him/her a long time to feel better | **.507** | .056 | .034 | .165 |
| 5 | 36. When my child is upset, his/her emotions feel overwhelming | **.502** | .062 | −.044 | .323 |

*Note.* DERS-P = Difficulties in Emotion Regulation Scale–Parent-report form. OF = original factor and indicates what factor each item loaded onto for the adolescent self-report version of the DERS (i.e., 1 = Nonacceptance; 2 = Goals; 3 = Impulse; 4 = Awareness; 5 = Strategies; 6 = Clarity). Boldface font indicates final factor solution.

participants and their primary caretaker). The two samples were obtained from two larger, multisite treatment studies, including a sample of middle school students with ADHD (see Evans et al., 2016; hereafter "middle school ADHD") and a sample of high school students with ADHD (see Owens et al., 2018, November; hereafter "high school ADHD").[2] In the middle school ADHD sample ($n$ = 78),[3] youth ranged in age from 10 to 14 years ($M$ = 12.12, $SD$ = 0.91) and were in Grades 6th through 10th ($M$ = 6.96; $SD$ = 0.78). They were predominantly boys (75.6%), non-Hispanic (98.7%), and White (89.7%). In the high school ADHD sample ($n$ = 206), youth ranged in age from 13 to 17 years ($M$ = 15.35; $SD$ = 0.85) and were in Grades 9th through 11th ($M$ = 9.75; $SD$ = 0.76) in high school. They were predominantly boys (78.2%), non-Hispanic (89.4%), and White (72.9%; black: 22.4%).

*Procedures.* Data were collected in the context of larger research studies, in the middle school ADHD sample, data were collected at posttreatment, in the high school ADHD

sample, data were collected at eligibility evaluations/pre-treatment. Data were combined across the two studies for purposes of confirming the factor structure of the DERS-P. Otherwise, data from the three available samples were analyzed separately to evaluate evidence of convergent, concurrent, and incremental validity of the DERS-P and its indicated factors.

*Measures.* In addition to the DERS-P, additional parent- and self-report measures were included to assess evidence of validity for the DERS-P. In some cases, because different measures of the same characteristic were used across studies (e.g., for reasons of age-appropriateness or respective research design), some measures were administered only to the online sample, middle school ADHD sample, or high school ADHD sample. These differences are noted below.

The Beck Youth Inventory II–Anxiety (BYI-II-Anxiety; Beck, Beck, Jolly, & Steer, 2005) is a 20-item self-report scale of anxiety symptoms administered to youth 7 to 18 years old. Each item is rated on a 4-point scale (0 = *never*; 3 = *always*), with higher scores indicative of greater anxiety. The BYI-II-Anxiety has demonstrated evidence of reliability and validity, including strong association with other self-report measures of anxiety (Beck et al., 2005). In the current research, the BYI-II-Anxiety demonstrated excellent internal consistency in the high school ADHD sample ($\alpha = .91$).

See above for item content and prior research on the psychometric properties of the DERS (Gratz & Roemer, 2004). The DERS has six subscales (as described above): Awareness (six items), Clarity (five items), Goals (five items), Impulse (six items), Nonacceptance (six items), and Strategies (eight items). Items are rated on a 5-point scale (1 = *almost never*; 5 = *almost always*), with higher scores indicative of greater ED. The DERS subscales demonstrated adequate internal consistency in the middle school ADHD and the high school ADHD samples across domains, respectively (i.e., Awareness: $\alpha = .61, .74$; Clarity: $\alpha = .62, .74$; Goals: $\alpha = .79, .82$; Impulse: $\alpha = .82, .84$; Nonacceptance: $\alpha = .86, .88$; and Strategies: $\alpha = .85, .91$).

The Disruptive Behavior Disorders Rating Scale (DBD; Pelham, Gnagy, Greenslade, & Milich, 1992) is a 45-item parent- and teacher-report scale of 18 symptoms of ADHD (including nine inattention [IA] items and nine hyperactivity/impulsivity [HI] items), 8 symptoms of oppositional defiant disorder (ODD), and 15 symptoms of conduct disorder (CD). Raters use a 4-point scale (0 = *not at all*; 3 = *very much*), to rate how well each item describes their child, with higher scores indicative of greater symptom severity. The DBD has demonstrated evidence of reliability and validity (cf. Massetti et al., 2003). In the current middle school ADHD sample, parent-report data were collected and internal consistency across all scales was acceptable (i.e., IA, $\alpha = .95$; HI, $\alpha = .90$; ODD, $\alpha = .92$; CD, $\alpha = .85$).

The Emotion Regulation Index for Children and Adolescents (ERICA; MacDermott, Gullone, Allen, King, & Tonge, 2010) is a 16-item self-report measure of emotion regulation. Each item is rated on a 5-point scale (1 = *strongly disagree*; 5 = *strongly agree*), with higher scores indicative of greater emotion regulation. The ERICA contains three subscales: Emotional Control (seven items reflective of a low threshold for emotional excitability, impatience, defiance, and undercontrolled anger), Emotional Self-Awareness (five items focused on adaptability or flexible emotion management and a generally euthymic emotional style to which the individual is able to quickly return), and Situational Responsiveness (four items related to exhibiting socially appropriate emotional responses in social situations). It has demonstrated adequate reliability and evidence of validity (Bunford et al., 2014; MacDermott et al., 2010). The ERICA evinced adequate internal consistency in both the middle school and the high school ADHD samples across domains, respectively (i.e., Emotional Control: $\alpha = .86, .78$; Self-Awareness: $\alpha = .62, .74$), except for the situational responsiveness subscale, which demonstrated poor internal consistency in the middle school sample ($\alpha = .21$), but adequate internal consistency in the high school ADHD sample ($\alpha = .64$).

The Multidimensional Anxiety Scale for Children (MASC; March, Parker, Sullivan, Stallings, & Conners, 1997) is a 39-item self-report scale of anxiety symptoms across four domains (i.e., harm avoidance, physical symptoms, separation, and social anxiety), administered to youth 8 to 19 years old. Each item is rated on a 4-point scale (0 = *never true about me*; 3 = *often true about me*), with higher scores indicative of greater anxiety. The MASC has demonstrated evidence of reliability and validity (Baldwin & Dadds, 2007; March et al., 1997). Item-level data were not available. We used the total anxiety *T* score from the middle school ADHD sample (based on age and sex norms from a norming sample of >9,000 adolescents).

The Index of Family Relations (IFR; Hudson, 1992) is a 25-item measure of the magnitude and severity of personal and social impairment in the context of family relations. Each item is rated on a 7-point scale (1 = *none of the time*; 7 = *all of the time*). Norms and standardization sample information reveal acceptable psychometric properties, and evidence of reliability and validity (Hudson, 1992). Self-report data were collected and the IFR demonstrated excellent internal consistency in the high school ADHD sample ($\alpha = .96$).

The Reynolds Adolescent Depression Scale–II (RADS-2; W. M. Reynolds, 2002) is a self-report scale of depressive symptoms in youth, across four domains: anhedonia/negative affect, dysphoric mood, negative self-evaluation, and somatic complaints. Each item is rated on a 4-point scale (1 = *almost never*; 4 = *most of the time*) with higher scores indicative of greater depression. The RADS-2 has demonstrated evidence

of internal consistency ($\alpha = 0.93$), reliability, and validity (W. M. Reynolds, 2002). The full, 30-item version was administered to the middle school ADHD sample (but did not have item-level data available) and the short form (10 items) was administered to the high school ADHD sample ($\alpha = .87$). In the current research, we used the total depression *T* score (based on age and sex norms from a norming sample of >2,500 youth).

The Strengths and Difficulties Questionnaire (SDQ; Goodman, 1997) is a 25-item scale that has been validated for international use, with standardized norms across age groups and genders (Goodman, 2001). The SDQ includes five subscales: conduct problems, hyperactivity/inattention, emotional symptoms, peer relationship problems, and prosocial behavior. Items are rated on a 3-point scale (0 = *not true*; 2 = *certainly true*). The SDQ has demonstrated good specificity and moderate sensitivity in community samples in the prediction of mental health disorders (Goodman, Ford, Simmons, Gatward, & Meltzer, 2000). The scale is designed for community samples of the general population. In MTurk sample, all five subscales demonstrated adequate internal consistency (i.e., conduct problems: $\alpha = .74$, hyperactivity/inattention: $\alpha = .79$, emotional symptoms: $\alpha = .78$, peer relationship problems: $\alpha = .66$, and prosocial behavior: $\alpha = .77$).

The Impairment Rating Scale (IRS; Fabiano et al., 2006) was developed to assess areas of daily functioning (academic, social) that can be particularly problematic for adolescents with behavioral problems (specifically for adolescents with ADHD; Evans et al., 2013). Seven items are scored on a continuous scale, converted to seven equal parts from (0 = *no problem*; 6 = *extreme problem*), with higher scores indicating greater impairment. The psychometric properties of the IRS have been measured in both clinic and community samples. The IRS demonstrates excellent temporal stability and evidence of convergent and discriminant validity. The IRS was administered to the MTurk and the middle school ADHD samples and the total item (i.e., "overall severity of your child's problem in functioning and overall need for treatment") was used in the current research.

*Data Analytic Plan.* To maximize sample size, evidence for validity of the factor structure obtained in Study 1 was examined in Study 2 using the combined middle school and high school ADHD samples. CFA was conducted in Mplus 7.4 (Muthén & Muthén, 2012) using the weighted least squares estimator that is robust to nonnormality of categorical data. We also referenced available modification indices (MI), which reflect the degree to which the chi-square value would improve should a model modification be made. Model fit was examined using the $\chi^2/df$ ratio, root mean square error of approximation (RMSEA), and the comparative fit index (CFI). Conventionally, a $\chi^2/df$ ratio of 5:1 (Hooper, Coughlan, & Mullen, 2008; Wheaton, Muthen, Alwin, & Summers, 1977), an RMSEA $\leqslant .10$, and a CFI $\geqslant .90$ indicate sufficient fit (Bentler, 1990) and $\chi^2/df$ ratio of 2, RMSEA $\leqslant .06$, and CFI $\geqslant .95$ indicate excellent fit (Hu & Bentler, 1999). Given that the DERS subscales are presumed to measure lower-order domains of a higher-order ED construct, the final four-factor correlated-traits CFA was compared with a higher-order CFA in which each factor from the correlated-traits model was specified as a lower-order factor and ED was specified as a higher-order factor. Both the correlated-traits and higher-order models were compared with a single-factor CFA in which the DERS was modeled as measuring a unidimensional ED construct.

Following CFA, we assessed for age and gender differences across DERS-P factors, using all three samples (i.e., online MTurk sample, middle school ADHD sample, and high school ADHD sample). We evaluated evidence of convergent, concurrent, and incremental validity of the DERS-P. To these ends, wherever possible, in the middle school ADHD sample and the high school ADHD sample, we examined bivariate correlations between the DERS-P and measures of the same phenomenon (i.e., emotion regulation) or measures of theoretically related characteristics (including those administered to the online sample, such as externalizing and internalizing disorder symptoms, social impairment, family impairment, and overall impairment). To assess for evidence of incremental validity, we examined if, in regression analyses predicting outcomes associated with ED, DERS-P factors accounted for variance beyond that accounted for by the adolescent self-report DERS subscales collected in the middle and high school ADHD samples.

## Results and Discussion

*Confirmatory Factor Analyses.* The four-factor structure obtained in Study 1 was assessed for cross-validation and model fit via CFA. Model fit was approaching, but not conformably achieving acceptable levels across fit indices, $\chi^2(371) = 1{,}454.90$, $p < .001$; RMSEA = .101 (90% CI = [.096, .107]); CFI = .90. MI indicated that DERS-P Item 22 ("When my child is upset, he/she knows that he/she can find a way to eventually feel better.") might best cross-load on Factor 3 (MI = 255.52) and Factor 4 (MI = 53.67). On examination of this item, which was indicated by the EFA to load onto Factor 1, its wording appeared inconsistent with the other items on that factor. Specifically, items on Factor 1 appeared to be related to catastrophizing secondary to negative emotions (e.g., *end up feeling very depressed, feel overwhelmed, become out of control*), whereas Item 22 appears to reflect a tempered approach to handling upset emotions, inconsistent with catastrophizing. As such, Item 22 was deleted and the CFA was conducted again.

**Table 2.** Means and Standard Deviations for the Factors of the DERS-P Across All Three Samples.

| | Sample | | | | |
|---|---|---|---|---|---|
| DERS-P Factors | MTurk total, *M* (*SD*) | MTurk no special needs, *M* (*SD*) | MTurk special needs, *M* (*SD*) | Middle school ADHD, *M* (*SD*) | High school ADHD, *M* (*SD*) |
| f1: Catastrophize | 1.98 (0.82) | 1.87 (0.73) | 2.71 (1.01) | 2.20 (0.79) | 2.21 (0.84) |
| f2: Negative Secondary | 1.89 (0.77) | 1.84 (0.72) | 2.25 (0.98) | 1.72 (0.71) | 1.86 (0.77) |
| f3: Attuned | 2.55 (0.88) | 2.48 (0.85) | 3.01 (0.94) | 2.88 (0.95) | 2.83 (0.92) |
| f4: Distracted | 2.57 (1.07) | 2.46 (0.99) | 3.43 (1.21) | 3.22 (0.98) | 3.32 (1.00) |

*Note.* DERS-P = Difficulties in Emotion Regulation Scale–Parent-report form. *t* tests comparing factor means across each group were performed. The middle school ADHD and high school ADHD sample did not differ significantly on DERS-P factors, f1: $t(282) = -0.12$, $p = .906$; f2: $t(282) = -1.45$, $p = .148$; f3: $t(282) = 0.44$, $p = .661$; f4: $t(282) = -0.79$, $p = .433$. MTurk total sample differed from the middle school ADHD sample across all factors, f1: $t(1,054) = 2.34$, $p = .020$; f2: $t(1,054) = -1.97$, $p = .049$; f3: $t(1,054) = 3.23$, $p = .001$; f4: $t(1,054) = 5.22$, $p < .001$. The MTurk special needs subsample also differed from the middle school ADHD sample on factors: f1: $t(184.92) = -4.21$, $p < .001$, Cohen's $d = 0.61$; f2: $t(189.30) = -4.06$, $p < .001$, Cohen's $d = 0.66$; but not f3 or f4. The MTurk sample differed from the high school ADHD sample across the following factors: f1: $t(1,182) = 3.77$, $p < .001$, Cohen's $d = 0.28$; f3: $t(1,182) = 4.15$, $p < .001$, Cohen's $d = 0.31$; and f4: $t(1,182) = 9.33$, $p < .001$, Cohen's $d = 0.72$; but not f2. The MTurk special needs subsample also differed from the high school ADHD sample on factors: f1: $t(199.83) = -4.83$, $p < .001$, Cohen's $d = 0.58$; f2: $t(189.61) = -3.93$, $p < .001$, Cohen's $d = 0.48$; but not f3 or f4. The MTurk no special needs subsample differed on all factors from the special needs subsample: f1: $t(128.49) = 9.00$, $p < .001$, Cohen's $d = 1.00$; f2: $t(128.70) = 4.60$, $p < .001$, Cohen's $d = 0.51$; f3: $t(976) = 5.95$, $p < .001$, Cohen's $d = 0.58$; f4: $t(133.07) = 8.93$, $p < .001$, Cohen's $d = 0.95$.

After deleting Item 22, model fit was marginally adequate considering all fit indices: $\chi^2(344) = 1{,}126.33$, $p < .001$; RMSEA = .089 (90% CI = [.084, .095]); CFI = .93. Additionally, all items had a standardized factor loading estimate of 0.45 or greater, further supporting the conclusion of sufficient fit. All four factors maintained acceptable internal consistency (i.e., Factor 1 (Catastrophize): $\alpha = .91$; Factor 2 (Negative Secondary): $\alpha = .89$; Factor 3 (Attuned): $\alpha = .89$; Factor 4 (Distracted): $\alpha = 84$). Factor 1 was correlated with Factor 2 ($r = .66$, $p < .001$), Factor 3 ($r = .14$, $p = .010$), and Factor 4 ($r = .68$, $p < .001$). Factor 2 was also correlated with Factor 4 ($r = .40$, $p < .001$), but not Factor 3 ($r = .09$, $p = .157$). Factor 3 was not correlated with Factor 4 ($r = .003$, $p = .954$).

The final four-factor correlated-traits model was compared with a higher-order CFA in which the four factors in the correlated-traits model (i.e., Catastrophize, Negative Secondary, Attuned, and Distracted) were specified as lower-order factors and ED was specified as a higher-order factor. A nonsignificant negative residual variance for the Catastrophize factor was fixed as zero. In the final model, model fit was marginally adequate across all indices: $\chi^2(347) = 1{,}036$, $p < .001$; RMSEA = .084 (90% CI = [.078, .090]); CFI = .93. All four lower-order factors loaded significantly on the higher-order ED factor, though the loading for Factor 3 (Attuned) was smaller in magnitude (loading = .12) than those for the other three factors (loadings $\geqslant .65$). Examination of the $R^2$ indicated that the ED factor did not account for significant variance in the Attuned factor. Results of a $\chi^2$ difference test for the weighted least squares estimator indicated that the higher-order model did not result in significantly worse fit, $\chi^2$ DIFFTEST = 4.40(3), $p = .22$.

Both the correlated-traits and higher-order models were compared with a single-factor model in which the DERS was modeled as measuring a unidimensional construct of ED. For the unidimensional model, model fit was poor across all fit indices: $\chi^2(350) = 4499.47$, $p < .001$; RMSEA = .204 (90% CI = [.199, .210]); CFI = .61. Results of $\chi^2$ difference tests for the weighted least squares estimator indicated that the single-factor model resulted in significantly worse fit compared with both the correlated-traits model, $\chi^2$ DIFFTEST = 720.20(6), $p < .001$, and the higher-order model, $\chi^2$ DIFFTEST = 604.22(3), $p < .001$.

*Evidence of Validity Across Factors.* Mean factor scores were created for the four DERS-P factors in Study 1 and the ADHD samples in Study 2 (see Table 2 for means and standard deviations). Notably, the more representative online MTurk sample demonstrated significantly lower parent-rated ED across factors than did the samples consisting of adolescents with ADHD (which did not significantly differ from each other). In addition to mean factor scores across samples, each factor was examined for differences, given adolescent age and gender. In the MTurk sample, age had a small negative association with ED on DERS-P Factor 1 (Catastrophize; $r = -.08$, $p = .008$) and Factor 2 (Negative Secondary; $r = -.06$, $p = .049$). Conversely, age was not associated with factors on the DERS-P in the ADHD samples. Gender differences emerged in the MTurk sample and in the high school ADHD sample. In the MTurk sample, boys ($M = 2.65$; $SD = 0.88$) had higher scores on Factor 3 (Attuned) than girls ($M = 2.43$; $SD = 0.86$), $t(976) = -3.93$, $p < .001$, Cohen's $d = 0.25$. In the high school sample, girls ($M = 2.55$; $SD = 0.92$) had higher scores on Factor 1 (Catastrophize) than boys ($M = 2.08$; $SD = 0.74$),

**Table 3.** Bivariate Correlations Indicating Convergent Validity and Consistency Across Parent- and Self-Report of Emotion Regulation.

| DERS-P factors | DERS-Nonacceptance | DERS-Goals | DERS-Impulse | DERS-Awareness | DERS-Strategies | DERS-Clarity | ERICA-Control | ERICA-SA | ERICA-SR |
|---|---|---|---|---|---|---|---|---|---|
| Middle school ADHD sample | | | | | | | | | |
| f1 | .23* | .06 | .19† | .17 | .25* | .06 | .05 | −.04 | −.05 |
| f2 | .03 | −.16 | −.01 | .001 | .06 | −.03 | .26* | −.07 | .02 |
| f3 | .24* | .03 | −.01 | .26* | .21† | .24* | −.17 | −.22† | −.22† |
| f4 | .19 | .22† | .18 | .13 | .27* | .04 | .14 | .08 | .01 |
| High school ADHD sample | | | | | | | | | |
| f1 | .09 | .06 | .20** | .13 | .21** | .13† | −.08 | −.24** | −.03 |
| f2 | .03 | .01 | .04 | .08 | .07 | .22** | .04 | −.19** | .04 |
| f3 | −.04 | −.03 | −.03 | .22** | .05 | .18** | −.08 | −.23** | −.13† |
| f4 | .05 | .10 | .10 | .07 | .07 | .05 | −.02 | −.10 | .16† |

*Note.* ADHD = attention deficit/hyperactivity disorder; DERS-P = Difficulties in Emotion Regulation Scale–Parent-report form; f1 = Catastrophize; f2 = Negative Secondary; f3 = Attuned; f4 = Distracted; ERICA = Emotion Regulation Index for Children and Adolescents; SA = Self-Awareness; SR = Situational Responsiveness. Shaded areas represent factors that share the most items from the adolescent DERS and the parent DERS-P.
†$p < .10$. *$p < .05$. **$p < .01$.

$t(172) = 3.25$, $p = .001$, Cohen's $d = 0.56$; and girls ($M = 3.71$; $SD = 0.94$) also had higher scores on Factor 4 (Distracted) than boys ($M = 3.21$; $SD = 0.95$), $t(172) = 2.89$, $p = .004$, Cohen's $d = 0.53$.

Next, we assessed for evidence of validity. First, we examined the degree to which the DERS-P exhibits evidence of concurrent validity as evaluated by bivariate correlations between the DERS-P and self-report emotion regulation on the DERS and ERICA (see Table 3). In general, the associations between the DERS-P and the DERS and ERICA were small to medium in magnitude (at marginal to significant levels). The pattern of associations between the DERS-P factors and DERS subscales was generally comparable across the two samples, with an association between DERS-P Factor 1 (Catastrophize) and the DERS Impulse subscale, the DERS-P Factor 3 (Attuned) and the DERS Awareness subscale, DERS-P Factor 1 (Catastrophize) and DERS Strategies subscale, and the DERS-P Factor 3 (Attuned) and the DERS Clarity subscale. The DERS-P Factors 2 (Negative secondary) and 4 (Distracted) exhibited the least number of associations with DERS subscales, with a single statistically significant association in the high school and the middle school samples, respectively.

The pattern of associations between the DERS-P and ERICA was not as comparable across the two samples. The DERS-P Factor 2 was positively associated with the ERICA Emotional Control subscale in the middle school ADHD sample, whereas DERS-P Factors 1, 2, and 3 were negatively associated with the ERICA Self-Awareness subscale in the high school ADHD sample.

We then examined the degree to which the DERS-P exhibits convergent validity as evaluated by bivariate correlations between the DERS-P and other conceptually related characteristics across all three samples (see Table 4, organized in a way to aid comparison across samples). Analyses of the online MTurk sample revealed that the DERS-P factors generally exhibited associations that are medium to high in magnitude in the expected direction with the characteristics of interest (i.e., conduct problems, hyperactivity, emotional problems, peer problems, and impairment). In the middle school ADHD sample, many associations were significant (and, when not significant, many were moderate in magnitude) and in the expected directions. In both the online MTurk and the middle school ADHD sample, parent-reported ADHD symptoms and conduct problems were associated with DERS-P factors. In the middle school ADHD sample, self-reported internalizing problems were not associated with DERS-P factors and in the high school ADHD sample, although these were associated, these associations were generally small in magnitude.

Finally, using the middle and the high school ADHD samples, we conducted hierarchical regression analyses to examine the degree to which the DERS-P factors exhibit incremental validity (i.e., variance accounted for) beyond that accounted for by factors of the self-report DERS in predicting parent and adolescent ratings of psychopathology (i.e., ADHD, ODD, CD, anxiety, and depression symptoms) and impairment (i.e., self-reported family impairment and parent-reported overall impairment). To this end, we entered self-report DERS subscales in Step 1 and DERS-P factors in Step 2. The entry of DERS subscales and DERS-P factors were organized in accordance with number of overlapping items and magnitude of bivariate association (e.g., DERS Subscales 3 and 5 were added to Step 1, followed by DERS-P Factor 1, as the items on these factors overlapped most). To minimize the number of tests, we examined

**Table 4.** Assessing Convergent Validity for the DERS-P Factors and Theoretically Associated Variables Across the Three Samples.

| | MTurk sample | | | | | |
|---|---|---|---|---|---|---|
| DERS-P factors | SDQ:emo | SDQ:peer | SDQ:hyp | SDQ:cndct | IRS:T | SDQ:pro |
| f1 | .62** | .46** | .59** | .63** | .62** | −.47** |
| f2 | .55** | .39** | .42** | .46** | .46** | −.28** |
| f3 | .39** | .35** | .46** | .43** | .43** | −.50** |
| f4 | .48** | .29** | .56** | .36** | .36** | −.27** |

| | Middle school ADHD sample | | | | | | |
|---|---|---|---|---|---|---|---|
| | MASC | RADS-2 | DBD:IA | DBD:HI | DBD:OD | DBD:CD | IRS:T |
| f1 | .04 | .06 | .29* | .40** | .65** | .35** | .55** |
| f2 | .20† | .09 | .33** | .40** | .34** | .22† | .42** |
| f3 | .10 | .07 | .20† | .15 | .19† | −.03 | .00 |
| f4 | .12 | .07 | .35** | .40** | .40** | .22† | .44** |

| | High school ADHD sample | | |
|---|---|---|---|
| | BYI-II-A | RADS-2-S | IFR |
| f1 | .29** | .23** | .12 |
| f2 | .19* | .18* | .09 |
| f3 | .12 | .12 | .21** |
| f4 | .08 | .13 | −.04 |

*Note*. ADHD = attention deficit/hyperactivity disorder; DERS-P = Difficulties in Emotion Regulation Scale–Parent-report form; f1 = Catastrophize; f2 = Negative Secondary; f3: Attuned; f4: Distracted; SDQ:emo = emotional symptoms (parent report); SDQ:peer = peer relationship problems (parent report); SDQ:hyp = hyperactivity/inattention (parent report); SDQ:cndct = conduct problems (parent report); SDQ:pro = prosocial behavior (parent report); IRS:T = overall impairment due to behavior (parent report); MASC = total anxiety score (adolescent report); RADS-2-(S)= total depression *t* score (adolescent report); DBD-IA = ADHD inattention (parent report); DBD:HI = ADHD hyperactivity/impulsivity (parent report); DBD:OD = ODD symptoms (parent report); DBD:CD = conduct disorder symptoms (parent report); BYI-II-A = Beck total anxiety score (adolescent report); IFR = total score from the Index of Family Relations (adolescent report).
†$p < .10$. *$p < .05$. **$p < .01$.

regressions only in cases wherein the DERS-P factor exhibited a statistically significant bivariate association with the outcome. In most analyses, DERS-P factors accounted for a significant amount of variance ($\Delta R^2$) beyond the contribution of the adolescent self-report DERS subscales ($R^2$; see Table 5). In some cases, the DERS-P factors accounted for a large percentage of the variance, for instance, DERS-P Factor 1 (Catastrophize) accounted for 36% of the additional variance in ODD symptoms beyond that of the Impulse and Strategies subscales of the DERS.

## General Discussion

Our goal in this research was to evaluate the psychometric properties of a newly created parent-report form of a measure of ED—the DERS-P form. Specifically, across two studies and three samples, we examined the (1) factor structure, (2) internal consistency, and evidence of (3) convergent, (4) concurrent, and (5) incremental validity of the DERS-P. Generally, the DERS-P demonstrated acceptable reliability and evidence of validity.

Regarding the factor structure of the DERS-P, a four-factor solution was obtained in Study 1 and confirmed with the ADHD samples in Study 2. The factors are Catastrophize, with items related to losing control over and feeling overwhelmed by negative emotions; Negative Secondary, with items related to a tendency to exhibit negative secondary emotions; Attuned, reflective of awareness or clarity of emotions; and Distracted, with items reflecting difficulty with concentrating or focusing, or completing tasks in the face of strong emotions. Each factor demonstrated acceptable internal consistency across studies. Interestingly, all factors correlated with each other (with at least a moderate-sized, $r > .4$ correlation) except for Factor 3 (Attuned), which correlated only and relatively weakly with Factor 1 (Catastrophize). These results are consistent those obtained with the self-report DERS indicating that all self-report subscales are generally correlated, but the Awareness subscale is, albeit meaningfully related to the other subscales (Fowler et al., 2014; Giromini et al., 2012; Neumann et al., 2010; Perez et al., 2012; Ritschel et al., 2015), typically exhibits a less strong correlation and sometimes exhibits no correlation. Although these results may be explained by there being commonalities among Factors 1, 2, and 4 (e.g., which seem to reflect the physiological and behavioral aspects of emotion regulation more than experiential or

**Table 5.** Assessing Incremental Validity of the DERS-P Relative to the Adolescent Self-Report DERS.

| | Middle school ADHD sample | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | MASC | | RADS-2 | | DBD:IA | | DBD:HI | | DBD:OD | | DBD:CD | | IRS:T | |
| DERS-P factors | Step 1[a] ($R^2$) | Step 2 ($\Delta R^2$) | Step 1 ($R^2$) | Step 2 ($\Delta R^2$) | Step 1 ($R^2$) | Step 2 ($\Delta R^2$) | Step 1 ($R^2$) | Step 2 ($\Delta R^2$) | Step 1 ($R^2$) | Step 2 ($\Delta R^2$) | Step 1 ($R^2$) | Step 2 ($\Delta R^2$) | Step 1 ($R^2$) | Step 2 ($\Delta R^2$) |
| f1 | — | — | — | — | .04 | .07* | .01 | .17** | .02 | .42** | .01 | .12** | — | — |
| f2 | — | — | — | — | .01 | .10** | .01 | .17** | .01 | .13** | .06 | .26** | .01 | .18** |
| f3 | — | — | — | — | — | — | — | — | — | — | — | — | — | — |
| f4 | — | — | — | — | .03 | .15** | .01 | .18** | .01 | .15** | — | — | .01 | .20** |

| High school ADHD sample | | | | | | |
|---|---|---|---|---|---|---|
| | BYI-II-A | | RADS-2-S | | IFR | |
| | Step 1 ($R^2$) | Step 2 ($\Delta R^2$) | Step 1 ($R^2$) | Step 2 ($\Delta R^2$) | Step 1 ($R^2$) | Step 2 ($\Delta R^2$) |
| f1 | .24** | .04* | .29** | .02* | — | — |
| f2 | .19** | .03* | .20** | .03* | — | — |
| f3 | — | — | — | — | .22** | .01 |
| f4 | — | — | — | — | — | — |

*Note.* ADHD = attention deficit/hyperactivity disorder; DERS-P = Difficulties in Emotion Regulation Scale–Parent-report form; f1 = Catastrophize; f2 = Negative secondary; f3: Attuned; f4: Distracted; IRS:T = overall impairment due to behavior; MASC = total anxiety score; RADS-2-(S)= total depression *t* score; DBD-IA = ADHD inattention; DBD:HI = ADHD hyperactivity/impulsivity; DBD:OD = ODD symptoms; DBD:CD = conduct disorder symptoms; BYI-II-A = Beck total anxiety score; IFR = total score from the Index of Family Relations.

[a]When f1 was entered on Step 2, f3 (Impulse) and f4 (Strategies) were entered for Step 1; when f2 was entered on Step 2, f1 (Nonacceptance) was entered for Step 1; when f3 was entered on Step 2, f4 (Awareness) and f6 (Clarity) were entered for Step 1; when f4 was entered on Step 2, f2 (Goals) was entered for Step 1.

$^{\dagger}p < .10$. $^{*}p < .05$. $^{**}p < .01$.

cognitive ones) relative to Factor 3 (e.g., which seems to highly reflect, if not exclusively, aspects of emotion regulation that are cognitive in nature) that account for such difference, it will be important to examine whether, similar to the self-report DERS (Bardeen et al., 2012; Bardeen et al., 2016; D. J. Lee et al., 2016), the reverse-coded items are driving the unexpected results.

In examining the factor structure of the DERS-P, we also considered a higher-order model and a unidimensional model. The poor overall fit of the unidimensional model, and the poor fit in comparison with the four-factor correlated-traits and higher-order models, indicate that the DERS-P is not measuring a unidimensional construct. Although comparison of the model fit indices for the higher-order and correlated-traits model suggests that the higher-order model provides a more parsimonious and superior fit to the data, the Attuned factor did not demonstrate expected relations to the higher-order ED factor (i.e., weak loading on the higher-order ED factor). This finding is inconsistent with the hypothesis that all four factors of the DERS-P in the correlated-traits model represent lower-order aspects of a higher-order ED construct but is consistent with the results of Bardeen and colleagues (e.g., Bardeen et al., 2012; Bardeen et al., 2016; D. J. Lee et al., 2016). It is unclear whether this finding is attributable to the above-noted differences between the underlying phenomena measured by Factors 1, 2, and 4 and the Attuned factor (i.e., physiological and behavioral aspects of emotion regulation vs. cognitive aspects) or are a result of item format, rather than content. As such, although our results provide some evidence of a higher-order structure for the DERS-P, taking into account the magnitude of the path coefficients from lower-order to higher-order factors (specifically, of Attuned to the higher-order factor), it will be important to examine the higher-order structure of this measure using forward-coded items as in Bardeen et al. (2016), to rule out the possibility that reverse-coded items are driving these effects.

Given prior findings indicating that, relative to typically developing children (Maedgen & Carlson, 2000; Melnick & Hinshaw, 2000) or a community sample of youth (Bunford et al., 2014; MacDermott et al., 2010), youth with ADHD exhibit greater ED than youth without ADHD, this finding was expected in the current study (i.e., comparing the more representative online sample with the two samples of adolescents with ADHD). Indeed, our online MTurk sample was rated as less emotionally dysregulated than both our middle school and high school ADHD samples on the Attuned, Catastrophize, and Distracted factors, further confirming the growing body of work indicating ED is at least an associated feature of ADHD, across development,

measurement methods, and samples. However, the online MTurk sample was rated as exhibiting greater ED on the Negative Secondary factor than the sample of middle school youth with ADHD and as exhibiting comparable ED on this factor to the high school youth with ADHD. As depicted in Table 1, the Negative Secondary factor includes items pertaining to feeling angry or embarrassed about being upset. These negative appraisals of oneself following feeling upset may indicate a degree of metacognitive ability, such as an awareness and reflection that many youths with ADHD do not exhibit (Schroeder & Kelley, 2009). Indeed, the younger adolescents with ADHD were rated significantly lower on this factor—both their youth and their ADHD symptoms may in combination make them particularly prone to lack of awareness, insight, and/or remorse.

There is literature indicating both the absence and the presence of age differences in ED or differential effects of age on the association between ED and other characteristics (Bunford et al., 2014; Bunford, Evans, et al., 2017; Bunford, Wymbs, et al., 2017; Graziano & Garcia, 2016; Zimmermann & Iwanski, 2014). Regarding findings suggesting the presence of age differences in ED, there are general age-related increases in adaptive emotion regulation, though relative to other developmental phases, middle adolescence is characterized by the smallest emotion regulation strategy repertoire (Zimmermann & Iwanski, 2014).

In this research, there was a small negative association between age and ED in our more representative online sample with the Catastrophize and Negative Secondary factors, but no such age-related decrease in ED was observed in our ADHD samples. These results may be mainly a factor of variability (i.e., wider age range in the representative sample) but may also suggest that although normatively, observable ED decreases with age, either parent ratings are not sensitive to this change or observable ED does not decrease with age in the case of adolescents with ADHD. As such, our findings raise questions about parents' ability to detect changes across development. Alternatively, or in addition, they may also suggest the gap in ED between youth with/without ADHD may widen from early adolescence into and through late adolescence, increasing potential for functional impairments, such as social dysfunction (Bunford et al., 2014; Bunford, Evans, Becker, et al., 2015). An exception to this may be aspects of ED measured on the Negative Secondary factor, as described above.

As with age, prior findings suggest both the absence and the presence of gender differences in ED (Bunford et al., 2014; Bunford, Evans, et al., 2017; Bunford, Wymbs, et al., 2017; Graziano & Garcia, 2016; Zimmermann & Iwanski, 2014). For example, there are gender differences with regard to primary motivation for use of emotion regulation strategies (i.e., downregulate negative vs. upregulate positive emotions; Mitchell et al., 2012) and to individuals' typically used emotion regulation strategies (Zimmermann & Iwanski, 2014). As in prior studies comparing men and women and boys and girls (with women/girls exhibiting greater ED on some of the factors, Bunford et al., 2014; Weinberg & Klonsky, 2009), gender differences emerged in the online MTurk sample and in our older ADHD sample (i.e., high school sample), but on different factors and in the opposite direction. In the MTurk sample, boys exhibited greater difficulties with awareness or clarity of emotions (Attuned), whereas in the high school sample, girls exhibited greater difficulty with losing control over and feeling overwhelmed by negative emotions (Catastrophize) as well as with concentrating, focusing, or completing tasks in the face of strong emotions (Distracted). It is of note that the MTurk sample was large and relatively well-balanced with regard to gender (51.2% boys), whereas the ADHD samples were predominantly boys (middle school ADHD: 75.6%; high school ADHD: 72.9%). The differences in attunement may be a product of a large enough sample of boys and girls wherein parents may attend to emotions in their children differently in a manner consistent with gender stereotypes (Chaplin, Cole, & Zahn-Waxler, 2005) and/or genuine differences in emotion expression across genders in adolescence (Chaplin & Aldao, 2013). The gender differences observed in the sample of older adolescents with ADHD should be considered with the population in mind. Boys are most frequently diagnosed with ADHD, yet available evidence indicates that girls with ADHD have at least as many, if not more, difficulties in peer relationships than do boys with ADHD (Mikami, 2010; Mikami & Hinshaw, 2003). It has been speculated that difficulties with the self-regulation of affect, motivation, and arousal are more likely to affect characteristic features of women/girls' interpersonal interactions, such as emotional intimacy, focus on reciprocity, and verbal give-and-take than characteristic features of men/boys' interactions (Mikami & Hinshaw, 2003). In addition, because ADHD is rarer in girls, a girl with ADHD may appear more deviant than a boy with ADHD (Mikami, 2010). Even more generally, in line with the gender paradox perspective positing that the less-prevalent gender with a given disorder shows greater impairment than the more prevalent gender with the disorder (Eme, 1992), girls with ADHD may warrant greater concern. These hypotheses may apply to ED, in association with ADHD as well insofar as our findings may reflect parents with girls with ADHD and ED perceiving their children as more impaired across some domains, including in processes related to the experience and expression of emotions, than parents with boys with ADHD and ED.

Regarding evidence of convergent validity between the DERS-P and self-report DERS, not surprisingly, the DERS-P factors were generally more strongly associated with DERS subscales when they included overlapping items (see Table 3). This differentiation in associations, wherein overlapping factors/subscales were more associated provides preliminary

evidence supporting the use of the DERS-P/DERS as parallel multi-informant reports of ED in future research. However, one exception to this was that the second DERS-P factor, Negative Secondary, was not associated with the DERS Nonacceptance subscale. Consistent with the above hypothesis speculating the Negative Secondary factor reflects lack of awareness, insight and/or remorse when upset, it may be that the lack of association between the DERS-P Negative Secondary factor and the DERS Nonacceptance subscale indicates that adolescents' anger or embarrassment over their own behaviors related to having been upset is a private event they do not admit to parents. As a result, parental perceptions of this remorse may be different from adolescents' self-reports.

Regarding the relationship between the DERS-P and self-report ERICA, there is some evidence of convergent validity in the high school ADHD sample. Greater parent-rated ED (as indexed by tendency to lose control over and feel overwhelmed by negative emotions; Catastrophize, to exhibit negative secondary emotions; Negative Secondary, and to have difficulty with awareness or clarity of emotions; Attuned) was associated with greater self-rated ED (as indexed by adaptability or flexible emotion management and a generally euthymic emotional style to which the individual is able to quickly return; ERICA Self-Awareness). However, the magnitude of correlations is small, possibly due to the different aspects and manifestations that the DERS and ERICA measure. Together, data on concurrent validity between these parent- and self-report rating scale measures of ED indicate such validity depends on the specific aspect of ED measured.

Regarding concurrent validity, in the more representative sample, DERS-P factors generally exhibited medium to high associations with characteristics of interest (i.e., conduct problems, hyperactivity/inattention, emotional and peer problems, prosocial behaviors, and functional impairment) in the expected direction. In the middle school ADHD sample, parent-rated ODD symptoms were associated with each of the four DERS-P factors (*r*s ranging from .19 to .65) and parent-rated H/I symptoms were associated with all DERS-P factors except Factor 3 (Attuned; all *r*s .40). These correlation coefficients were large, but not so large as to indicate isomorphism. There was an interesting pattern of results in the high school ADHD sample, where both domains of self-rated internalizing pathology—anxiety and depression—were associated (with a small magnitude) with parent-rated tendency to lose control over and feel overwhelmed by negative emotions (Catastrophize) and to exhibit negative secondary emotions (Negative Secondary), but not the other two DERS-P factors. Of note, the weak relationships between anxiety and depression with the Negative Secondary factor are also consistent with the explanatory hypothesis speculating this factor is related to the tendency for self-reflection and remorse over one's behavior exhibited when upset. Self-reported severity of personal and social impairment in the context of family relations was associated with parent-rated awareness or clarity of emotions (Attuned), but not the other three DERS-P factors.

Finally, in the majority of regression analyses, DERS-P factors accounted for a significant amount of variance beyond the contribution of the self-report DERS in support of incremental validity of the DERS-P. This was true for analyses with other parent-report measures as the dependent variable as well as with self-report measures as the dependent variables, indicating that the additional variance accounted for was not solely a function of source. As such, using the DERS-P in isolation or in addition to the DERS is likely to provide important contextual information. Most adolescents are living with and relying on their parents and it is likely that a child's ED has an effect on family functioning (and vice versa; Bunford, Evans, & Wymbs, 2015). Understanding parents' perspective of adolescent ED may be important, particularly considering that we observed that parent-reported ED accounts for considerable variance in ODD symptoms and other externalizing behaviors. Parents' view of their child's ED may be helpful in understanding the family context and how the family functions in reaction to the adolescent's ED, ideally to inform interventions focused on the role of child ED in family dysfunction and the role of a potentially emotionally invalidating environment for a child exhibiting ED (Bunford, Evans, & Wymbs, 2015).

### *Limitations and Directions for Future Research*

There are several limitations to this research that are important to note. First, in our CFA, model fit was marginally adequate, indicating it will be important to identify changes necessary to improve model fit, for example, as already and extensively discussed above by forward-coding all reverse-coded items (also see below). To improve model fit, we removed Item 22 arguing that although the EFA indicated for it to load onto Factor 1, its wording is inconsistent with the rest of the items. As already discussed, the issue may not be related to item content, but to it being reverse-coded and the forward-coded version may be more consistent with the "catastrophizing" content of the first factor.

Second, it could be expected for parent-rated ED to be more strongly correlated with other parent-rated characteristics than self-rated ED simply due to rater/within-source bias. However, we remain optimistic about the promise of our findings of associations between parent-report and teen self-report, somewhat mollifying the concern of method bias. Nevertheless, future research may include measures of other characteristics of interest and measure them via self-report or ratings obtained from peers or teachers and expand to methodologies assessing other correlates or indices of ED—such as behavioral, cognitive, physiological, and

neural assessment to within-level-of-measurement bias (Morris & Cuthbert, 2012).

Third, data were collected at posttreatment in the middle school ADHD sample, and there is reason to believe that at least one of the program components, the Interpersonal Skills Group (Evans, Schultz, DeMars, & Davis, 2011), could have some beneficial effects on ED, even in the format in which it was delivered to the middle school ADHD sample (Bunford & Evans, 2017). Data having been collected at posttreatment may thus contribute to the smaller magnitudes of the observed associations between parent-reported ED and other indices of affective/behavioral functioning (also collected at posttreatment), due to broad treatment gains (Evans et al., 2016). However, these data likely did not affect evidence of construct validity, as what is important in this regard is that item variability and factor loading are similar regardless of intervention status or when data are collected. Relatedly, comparison of ED scores across the three samples should be made with caution as the MTurk and high school ADHD samples did not receive intervention from the researchers at the time of data collection (but, as noted, the middle school ADHD sample did).

Fourth, a multitude of tests were performed, raising possibility of Type I error. Our results are encouraging by a generally consistent pattern of results across multiple samples—especially considering the magnitude of effects, which has been argued to be as, if not more, meaningful and appropriate to interpret than *p* values (Baker, 2016; Wasserstein, 2016).

Fifth, there are some limitations inherent to the original, self-report measure that we modified to create the parent-report form. Specifically, some of the limitations regarding the factor structure of the self-report DERS may also be applicable to the parent-report version; the problems that the retained reverse-coded items have been shown to cause with the self-report measure, may be plaguing the herein developed parent-report measure as well. All six of the Factor 3 items are reverse-coded, and, consistent with Bardeen et al. (2016), all of these items are made up of reworded self-report Awareness and Clarity items, potentially consistent with a method effect. A solution to explore in future research and thus a primary task in further validating this parallel, parent-report form is to evaluate factor structure with all reverse-coded items forward-coded as in Bardeen et al. (2016). Another pertinent consideration is that the item content of the DERS is exclusively focused on dysregulation of negative emotions, despite data underscoring the importance and impairing potential of abnormalities in processing or regulating positive emotions (Bunford, Evans, & Wymbs, 2015; Bunford, Kujawa, et al., 2017; Okado & Mueller, 2016; Sjöwall et al., 2015).

Sixth, we included youth with ages spanning late childhood through late adolescence. Although we statistically evaluated the effects of age in our analyses, longitudinal designs are needed to clarify the developmental course of the effects we observed or did not observe. Related, we did not assess and thus could not statistically account for pubertal status, which might influence emotional processing (Silk et al., 2003).

Finally, caution should be taken when generalizing the results of our study. For instance, although our CFA demonstrated marginally adequate model fit for factor structure, it did not meet strict benchmarks for excellent fit. Our use of an ADHD-only sample to confirm factor structure may have introduced some range restriction on variance inherent in ADHD adolescent samples, and future cross-validation examinations should be performed to provide further supporting evidence. Moreover, although the MTurk worker population is generally more representative than college students or other community samples, it is nevertheless a *non*probability sample (Chandler & Shapiro, 2016). MTurk workers better represent Internet users than non-Internet users and are generally younger and better educated than the general population (Chandler & Shapiro, 2016). Related, although we employed quality control checks and given the number of consented participants relative to the number who completed the survey, confidence can be placed in sufficiently decreased likelihood of robots completing the online questionnaire, to ensure that not only "super-Turkers" (Bohannon, 2011) who often perform better than face-to-face participants (e.g., Hauser & Schwarz, 2016) but a representative sample of parents complete our survey, we chose a somewhat lower approval rating threshold (85%) than is typical (i.e., 95%; Bohannon, 2011). These considerations may limit the generality of our findings.

Although our primary conclusions regarding the middle school sample do not involve these two subscales, caution should also be taken when interpreting any of the findings obtained with the middle school ADHD sample involving the self-report Awareness and Clarity subscales as those exhibited questionable internal consistency in that sample.

## Conclusion

In conclusion, the current results provide support for the reliability and evidence of validity of the DERS-P for parent-report assessment of ED. Our findings demonstrate that although it will be a crucial next step in this line of research to examine whether some of the recently surfaced potential problems with the self-report DERS also apply to this parent-report version (i.e., related to the Awareness subscale and reverse-coded items; Bardeen et al., 2016), our results also suggest that the DERS-P is otherwise a psychometrically sound tool for the measurement of some parent-reported aspects of deficits in emotion regulatory capacities in adolescents aged between 11 and 17 years, including losing control over and feeling overwhelmed by

negative emotions; a tendency to exhibit negative secondary emotions; lack of awareness or clarity of emotions; and difficulty with concentrating or focusing, or completing tasks in the face of strong negative emotions. Given the lack of an otherwise suitable parent-report measure of adolescent ED, the present study makes an important contribution to the developmental and assessment literatures.

## Authors' Note

Portions of this study were presented at the 2017 biennial International Society for Research in Child & Adolescent Psychopathology convention in Amsterdam, the Netherlands.

## Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

## Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This study was conducted using data collected as part of two larger investigations supported by Research Grant R01 MH082864 funded by the National Institutes for Mental Health and awarded to SWE and JML and by R305A140356 funded by the Institute of Educational Sciences and awarded to SWE and GJD. During the preparation of this article, Nóra Bunford was funded by the MTA Premium Postdoctoral Fellowship (#PREMIUM-2017-67) and by the New National Excellence Program 2017/2018 (ÚNKP-17-4-ELTE/12423/11).

## Notes

1. Parent-report options are available to assess aspects of ED or characteristics that closely approximate it on many broadband symptom scales (e.g., Behavior Assessment System for Children [BASC; C. R. Reynolds & Kamphaus, 1992], Behavior Rating Inventory of Executive Function [BRIEF; Gioia, Isquith, Guy, & Kenworthy, 2000], Child Behavior Checklist [CBCL; Achenbach, 1991]), but there are no comprehensive measures of emotion regulation by parent report and the "emotion regulation" subscales on these measures do not have specific validity evidence supporting the "ED" component. Adding a valid parent-report measure of ED to the available assessments will enhance our ability to accurately measure the construct and partially address some of the limitations described above.
2. See Evans et al. (2016) and Owens et al. (2018, November and IES grant #R305A140356) for detailed information on design and participants.
3. In the middle school sample, more than 300 participants participated across 3 years; however, our sample size is smaller because the DERS-P was only administered to the third cohort (at Time-point 5, i.e., posttreatment).

## ORCID iD

Nóra Bunford https://orcid.org/0000-0002-9645-604X

## References

Achenbach, T. M. (1991). *Manual for the Child Behavior Checklist/4-18 and 1991 profile*. Burlington: University of Vermont.

Adrian, M., Zeman, J., Erdley, C., Lisa, L., Homan, K., & Sim, L. (2009). Social contextual links to emotion regulation in an adolescent psychiatric inpatient population: Do gender and symptomatology matter? *Journal of Child Psychology and Psychiatry*, *50*, 1428-1436. doi:10.1111/j.1469-7610.2009.02162.x

Baker, M. (2016). Statisticians issue warning over misuse of *P* values. *Nature*, *531*, 151. doi:10.1038/nature.2016.19503

Baldwin, J. S., & Dadds, M. R. (2007). Reliability and validity of parent and child versions of the Multidimensional Anxiety Scale for children in community samples. *Journal of the American Academy of Child & Adolescent Psychiatry*, *46*, 252-260. doi:10.1097/01.chi.0000246065.93200.a1

Bardeen, J. R., Fergus, T. A., Hannan, S. M., & Orcutt, H. K. (2016). Addressing psychometric limitations of the Difficulties in Emotion Regulation Scale through item modification. *Journal of Personality Assessment*, *98*, 298-309. doi:10.1080/00223891.2015.1091774

Bardeen, J. R., Fergus, T. A., & Orcutt, H. K. (2012). An examination of the latent structure of the Difficulties in Emotion Regulation Scale. *Journal of Psychopathology and Behavioral Assessment*, *34*, 382-392. doi:10.1007/s10862-012-9280-y

Barkley, R. A. (2010). Deficient emotional self-regulation: A core component of attention-deficit/hyperactivity disorder. *Journal of ADHD & Related Disorders*, *1*(2), 5-37.

Beck, J. S., Beck, A. T., Jolly, J. B., & Steer, R. A. (2005). *Manual for the Beck Youth Inventories of Emotional and Social Impairment* (2nd ed.). San Antonio, TX: Psychological Corporation.

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, *107*, 238-246. doi:10.1037/0033-2909.107.2.238

Bohannon, J. (2011). Social science for pennies. *Science*, *334*, 307. doi:10.1126/science.334.6054.307

Brady, C. E., Evans, S. W., Berlin, K. S., Bunford, N., & Kern, L. (2012). Evaluating school impairment with adolescents using the Classroom Performance Survey. *School Psychology Review*, *41*, 429-446.

Bunford, N., & Evans, S. W. (2017). Emotion regulation and social functioning in adolescence: Conceptualization and treatment. In J. R. Harrison, B. K. Schultz & S. W. Evans (Eds.), *School mental health services for adolescents* (1st ed., pp. 161-181). New York, NY: Oxford University Press.

Bunford, N., Evans, S. W., Becker, S. P., & Langberg, J. M. (2015). Attention-deficit/hyperactivity disorder and social skills in youth: A moderated mediation model of emotion dysregulation and depression. *Journal of Abnormal Child Psychology*, *43*, 283-296. doi:10.1007/s10802-014-9909-2

Bunford, N., Evans, S. W., & Langberg, J. M. (2014). Emotion dysregulation is associated with social impairment among young adolescents with ADHD. *Journal of Attention Disorders*, *22*, 62-82. doi:10.1177/1087054714527793

Bunford, N., Evans, S. W., & Wymbs, F. (2015). ADHD and emotion dysregulation among children and adolescents.

*Clinical Child and Family Psychology Review*, *18*, 185-217. doi:10.1007/s10567-015-0187-5

Bunford, N., Evans, S. W., Zoccola, P. M., Owens, J. S., Flory, K., & Spiel, C. F. (2017). Correspondence between heart rate variability and emotion dysregulation in children, including children with ADHD. *Journal of Abnormal Child Psychology*, *45*, 1325-1337. doi:10.1007/s10802-016-0257-2

Bunford, N., Kujawa, A., Fitzgerald, K. D., Swain, J. E., Hanna, G. L., Koschmann, E., . . . Phan, K. L. (2017). Neural reactivity to angry faces predicts treatment response in pediatric anxiety. *Journal of Abnormal Child Psychology*, *45*, 385-395. doi:10.1007/s10802-016-0168-2

Bunford, N., Wymbs, B. T., Dawson, A. E., & Shorey, R. C. (2017). Childhood maltreatment, emotional lability, and alcohol problems in young adults at-risk for ADHD: Testing moderation and moderated moderation. *Journal of Psychoactive Drugs*, *49*, 316-325. doi:10.1080/02791072.2017.1325031

Chandler, J., & Shapiro, D. N. (2016). Conducting clinical research using crowdsourced convenience samples. *Annual Review of Clinical Psychology*, *12*, 53-81. doi:10.1146/annurev-clinpsy-021815-093623

Chaplin, T. M., & Aldao, A. (2013). Gender differences in emotion expression in children: A meta-analytic review. *Psychological Bulletin*, *139*, 735-765. doi:10.1037/a0030737

Chaplin, T. M., Cole, P. M., & Zahn-Waxler, C. (2005). Parental socialization of emotion expression: Gender differences and relations to child adjustment. *Emotion*, *5*, 80-88. doi:10.1037/1528-3542.5.1.80

Costello, A., & Osborne, J. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment Research & Evaluation*, *10*(7), 1-9.

Coutinho, J., Ribeiro, E., Ferreirinha, R., & Dias, P. (2010). The Portuguese version of the Difficulties in Emotion Regulation Scale and its relationship with psychopathological symptoms. *Revista de Psiquiatria Clínica*, *37*, 145-151. doi:10.1590/S0101-60832010000400001

De Los Reyes, A., Augenstein, T. M., Wang, M., Thomas, S. A., Drabick, D. A. G., Burgers, D. E., & Rabinowitz, J. (2015). The validity of the multi-informant approach to assessing child and adolescent mental health. *Psychological Bulletin*, *141*, 858-900. doi:10.1037/a0038498

DeVellis, R. F. (2016). *Scale development: Theory and applications*. Thousand Oaks, CA: Sage.

Ehring, T., Fischer, S., Schnülle, J., Bösterling, A., & Tuschen-Caffier, B. (2008). Characteristics of emotion regulation in recovered depressed versus never depressed individuals. *Personality and Individual Differences*, *44*, 1574-1584. doi:10.1016/j.paid.2008.01.013

Eme, R. F. (1992). Selective female affliction in the developmental disorders of childhood: A literature review. *Journal of Clinical Child & Adolescent Psychology*, *21*, 354-364. doi:10.1207/s15374424jccp2104_5

Evans, S. W., Brady, C. E., Harrison, J. R., Bunford, N., Kern, L., State, T., & Andrews, C. (2013). Measuring ADHD and ODD symptoms and impairment using high school teachers' ratings. *Journal of Clinical Child & Adolescent Psychology*, *42*, 197-207. doi:10.1080/15374416.2012.738456

Evans, S. W., Langberg, J. M., Schultz, B. K., Vaughn, A., Altaye, M., Marshall, S. A., & Zoromski, A. K. (2016). Evaluation of a school-based treatment program for young adolescents with ADHD. *Journal of Consulting and Clinical Psychology*, *84*(1), 15-30. doi:10.1037/ccp0000057

Evans, S. W., Schultz, B. K., DeMars, C. E., & Davis, H. (2011). Effectiveness of the challenging horizons after-school program for young adolescents with ADHD. *Behavior Therapy*, *42*, 462-474. doi:10.1016/j.beth.2010.11.008

Fabiano, G. A., Pelham, W., Waschbusch, D. A., Gnagy, E. M., Lahey, B. B., & Chronis, A. M. (2006). A practical measure of impairment: Psychometric properties of the impairment rating scale in samples of children with attention deficit hyperactivity disorder and two school-based samples. *Journal of Clinical Child & Adolescent Psychology*, *35*, 369-385.

Fowler, J. C., Charak, R., Elhai, J. D., Allen, J. G., Frueh, B. C., & Oldham, J. M. (2014). Construct validity and factor structure of the Difficulties in Emotion Regulation Scale among adults with severe mental illness. *Journal of Psychiatric Research*, *58*, 175-180. doi:10.1016/j.jpsychires.2014.07.029

Fox, H. C., Axelrod, S. R., Paliwal, P., Sleeper, J., & Sinha, R. (2007). Difficulties in emotion regulation and impulse control during cocaine abstinence. *Drug and Alcohol Dependence*, *89*, 298-301. doi:10.1016/j.drugalcdep.2006.12.026

Fox, H. C., Hong, K. A., & Sinha, R. (2008). Difficulties in emotion regulation and impulse control in recently abstinent alcoholics compared with social drinkers. *Addictive Behaviors*, *33*, 388-394. doi:10.1016/j.addbeh.2007.10.002

Gao, Y., & Zhang, W. (2016). Confirmatory factor analyses of self- and parent-report inventory of callous-unemotional traits in 8- to 10-year-olds. *Journal of Psychopathology and Behavioral Assessment*, *38*, 331-340. doi:10.1007/s10862-015-9527-5

Gioia, G. A., Isquith, P. K., Guy, S. C., & Kenworthy, L. (2000). *Behavior Rating Inventory of Executive Function*. Odessa, FL: Psychological Assessment Resources.

Giromini, L., Velotti, P., De Campora, G., Bonalume, L., & Cesare Zavattini, G. (2012). Cultural adaptation of the Difficulties in Emotion Regulation Scale: Reliability and validity of an Italian version. *Journal of Clinical Psychology*, *68*, 989-1007. doi:10.1002/jclp.21876

Goodman, R. (1997). The Strengths and Difficulties Questionnaire: A research note. *Journal of Child Psychology and Psychiatry*, *38*, 581-586.

Goodman, R. (2001). Psychometric properties of the Strengths and Difficulties Questionnaire. *Journal of the American Academy of Child & Adolescent Psychiatry*, *40*, 1337-1345. doi:10.1097/00004583-200111000-00015

Goodman, R., Ford, T., Simmons, H., Gatward, R., & Meltzer, H. (2000). Using the Strengths and Difficulties Questionnaire (SDQ) to screen for child psychiatric disorders in a community sample. *International Review of Psychiatry (Abingdon, England)*, *15*, 166-172. doi:10.1080/09540260210000461 28

Gratz, K. L., & Roemer, L. (2004). Multidimensional assessment of emotion regulation and dysregulation: Development, factor structure, and initial validation of the Difficulties in Emotion Regulation Scale. *Journal of Psychopathology and Behavioral Assessment*, *26*, 41-54). doi:10.1023/B:JOBA.0000007455.08539.94

Graziano, P., & Garcia, A. (2016). Attention-deficit hyperactivity disorder and children's emotion dysregulation: A meta-analysis. *Clinical Psychology Review*, *46*, 106-123. doi:10.1016/j.cpr.2016.04.011

Hauser, D. J., & Schwarz, N. (2016). Attentive turkers: MTurk participants perform better on online attention checks than do subject pool participants. *Behavior Research Methods*, *48*, 400-407. doi:10.3758/s13428-015-0578-z

Hervás, G. (2011). Psicopatología de la regulación emocional: El papel de los déficit emocionales en los trastornos clínicos [Psychopathology of emotional regulation: The role of emotional deficit in clinical disorders]. *Behavioral Psychology/Psicologia Conductual*, *19*, 347-372.

Hooper, D., Coughlan, J., & Mullen, M. R. (2008). Structural equation modelling: Guidelines for determining model fit. *Electronic Journal of Business Research Methods*, *6*, 53-60.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*, 1-55. doi:10.1080/10705519909540118

Hudson, W. W. (1992). *The WALMYR assessment scale scoring manual*. Tallahassee, FL: WALMYR.

Lee, C., Milich, R., Lorch, E. P., Flory, K., Owens, J. S., Lamont, A. E., & Evans, S. W. (2017). Forming first impressions of children: The role of attention-deficit/hyperactivity disorder symptoms and emotion dysregulation. *Journal of Child Psychology and Psychiatry*. doi:10.1111/jcpp.12835

Lee, D. J., Witte, T. K., Bardeen, J. R., Davis, M. T., & Weathers, F. W. (2016). A factor analytic evaluation of the Difficulties in Emotion Regulation Scale. *Journal of Clinical Psychology*, *72*, 933-946. doi:10.1002/jclp.22297

MacDermott, S. T., Gullone, E., Allen, J. S., King, N. J., & Tonge, B. (2010). The Emotion Regulation Index for Children and Adolescents (ERICA): A psychometric investigation. *Journal of Psychopathology and Behavioral Assessment*, *32*, 301-314.

Maedgen, J. W., & Carlson, C. L. (2000). Social functioning and emotional regulation in the attention deficit hyperactivity disorder subtypes. *Journal of Clinical Child Psychology*, *29*, 30-42. doi:10.1207/S15374424jccp2901_4

March, J. S., Parker, J. D., Sullivan, K., Stallings, P., & Conners, C. K. (1997). The Multidimensional Anxiety Scale for Children (MASC): Factor structure, reliability, and validity. *Journal of the American Academy of Child & Adolescent Psychiatry*, *36*, 554-565.

Mash, E. J., & Hunsley, J. (2005). Evidence-based assessment of child and adolescent disorders: Issues and challenges. *Journal of Clinical Child & Adolescent Psychology*, *34*, 362-379. doi:10.1207/s15374424jccp3403_1

Massetti, G. M., Pelham, W., Chacko, A., Walker, K., Arnold, F., Keenan, J., . . . Burrows-MacLean, L. (2003, November). *Situational variability of ADHD, ODD, and CD: Psychometric properties of the DBD interview and rating scale*. Poster presented at the Association for Behavioral and Cognitive Therapies, Boston, MA.

Mazefsky, C. A., Herrington, J., Siegel, M., Scarpa, A., Maddox, B. B., Scahill, L., & White, S. W. (2013). The role of emotion regulation in autism spectrum disorder. *Journal of the American Academy of Child & Adolescent Psychiatry*, *52*, 679-688.

McDermott, M. J., Tull, M. T., Gratz, K. L., Daughters, S. B., & Lejuez, C. W. (2009). The role of anxiety sensitivity and difficulties in emotion regulation in posttraumatic stress disorder among crack/cocaine dependent patients in residential substance abuse treatment. *Journal of Anxiety Disorders*, *23*, 591-599. doi:10.1016/j.janxdis.2009.01.006

McQuade, J. D., Penzel, T. E., Silk, J. S., & Lee, K. H. (2016). Parasympathetic nervous system reactivity moderates associations between children's executive functioning and social and academic competence. *Journal of Abnormal Child Psychology*, *45*, 1355-1367. doi:10.1007/s10802-016-0246-5

Melnick, S. M., & Hinshaw, S. P. (2000). Emotion regulation and parenting in AD/HD and comparison boys: Linkages with social behaviors and peer preference. *Journal of Abnormal Child Psychology*, *28*, 73-86.

Merwood, A., Chen, W., Rijsdijk, F., Skirrow, C., Larsson, H., Thapar, A., . . . Asherson, P. (2014). Genetic associations between the symptoms of attention-deficit/hyperactivity disorder and emotional lability in child and adolescent twins. *Journal of the American Academy of Child & Adolescent Psychiatry*, *53*, 209-220.e4. doi:10.1016/j.jaac.2013.11.006

Mikami, A. Y. (2010). The importance of friendship for youth with attention-deficit/hyperactivity disorder. *Clinical Child and Family Psychology Review*, *13*, 181-198.

Mikami, A. Y., & Hinshaw, S. P. (2003). Buffers of peer rejection among girls with and without ADHD: The role of popularity with adults and goal-directed solitary play. *Journal of Abnormal Child Psychology*, *31*, 381-397. doi:10.1023/A:1023839517978

Mitchell, J. T., Van Voorhees, E. E., Dennis, M. F., McClernon, F. J., Calhoun, P. S., Kollins, S. H., & Beckham, J. C. (2012). Assessing the role of attention-deficit/hyperactivity disorder symptoms in smokers with and without posttraumatic stress disorder. *Nicotine & Tobacco Research*, *14*, 986-992.

Mitsopoulou, E., Kafetsios, K., Karademas, E., Papastefanakis, E., & Simos, P. G. (2013). The Greek version of the Difficulties in Emotion Regulation Scale: Testing the factor structure, reliability and validity in an adult community sample. *Journal of Psychopathology and Behavioral Assessment*, *35*, 123-131. doi:10.1007/s10862-012-9321-6

Morris, S. E., & Cuthbert, B. N. (2012). Research domain criteria: Cognitive systems, neural circuits, and dimensions of behavior. *Dialogues in Clinical Neuroscience*, *14*, 29-37.

Musser, E. D., Backs, R. W., Schmitt, C. F., Ablow, J. C., Measelle, J. R., & Nigg, J. T. (2011). Emotion regulation via the autonomic nervous system in children with attention-deficit/hyperactivity disorder (ADHD). *Journal of Abnormal Child Psychology*, *39*, 841-852.

Musser, E. D., Galloway-Long, H. S., Frick, P. J., & Nigg, J. T. (2013). Emotion regulation and heterogeneity in attention-deficit/hyperactivity disorder. *Journal of the American Academy of Child & Adolescent Psychiatry*, *52*, 163-171.

Muthén, L., & Muthén, B. (2012). *Mplus user's guide* (Version 7.4). Los Angeles, CA: Muthén & Muthén.

Neumann, A., van Lier, P. A. C., Gratz, K. L., & Koot, H. M. (2010). Multidimensional assessment of emotion regulation difficulties in adolescents using the Difficulties in Emotion Regulation Scale. *Assessment*, *17*, 138-149. doi:10.1177/1073191109349579

Nunnally, J. C. (1978). *Psychometric theory*. New York, NY: McGraw-Hill.

Okado, I., & Mueller, C. W. (2016). The relationship between child-reported positive affect and parent-reported emotional and behavioral problems in ADHD youth. *Journal of Child and Family Studies*, *10*, 2954-2965. doi:10.1007/s10826-016-0458-x

Owens, J. S., Evans, S. W., DuPaul, G. J., Benson, K., Belk, G., Puzino, K., Monopoli, J. & Wertz, E. (2018, November). The Bridges to Educational Success for Students (BEST) Program: Method, Design, and Intervention Integrity. In S. W. Evans (Chair) (Ed.), School-based treatment of high school students with ADHD: Results of a large randomized controlled trial. Symposium conducted at the annual meeting of the Association for the Advancement of Cognitive and Behavioral Therapy, Washington D.C.

Pelham, W., Gnagy, E. M., Greenslade, K. E., & Milich, R. (1992). Teacher ratings of *DSM-III-R* symptoms for the disruptive behavior disorders. *Journal of the American Academy of Child & Adolescent Psychiatry*, *31*, 210-218. doi:10.1097/00004583-199203000-00006

Perez, J., Venta, A., Garnaat, S., & Sharp, C. (2012). The Difficulties in Emotion Regulation Scale: Factor structure and association with nonsuicidal self-injury in adolescent inpatients. *Journal of Psychopathology and Behavioral Assessment*, *34*, 393-404. doi:10.1007/s10862-012-9292-7

Reynolds, C. R., & Kamphaus, R. W. (1992). *Behavior assessment system for children*. Circle Pinos, MN: American Guidance Service.

Reynolds, W. M. (2002). *Reynolds Adolescent Depression Scale* (2nd ed.). Hoboken, NJ: Wiley.

Ritschel, L. A., Tone, E. B., Schoemann, A. M., & Lim, N. E. (2015). Psychometric properties of the Difficulties in Emotion Regulation Scale across demographic groups. *Psychological Assessment*, *27*, 944-954. doi:10.1037/pas0000099

Rorhbeck, C. A., Azar, S. T., & Wagner, P. E. (1991). Child Self-Control Rating Scale: Validation of a child self-report measure. *Journal of Clinical Child Psychology*, *20*, 179-183. doi:10.1207/s15374424jccp2002_9

Rosen, P. J., & Epstein, J. N. (2010). A pilot study of ecological momentary assessment of emotion dysregulation in children. *Journal of ADHD & Related Disorders*, *1*(4), 39-52.

Rosen, P. J., Epstein, J. N., & Van Orden, G. (2013). I know it when I quantify it: Ecological momentary assessment and recurrence quantification analysis of emotion dysregulation in children with ADHD. *ADHD Attention Deficit and Hyperactivity Disorders*, *5*, 283-294. doi:10.1007/s12402-013-0101-2

Salters-Pedneault, K., Roemer, L., Tull, M. T., Rucker, L., & Mennin, D. S. (2006). Evidence of broad deficits in emotion regulation associated with chronic worry and generalized anxiety disorder. *Cognitive Therapy and Research*, *30*, 469-480. doi:10.1007/s10608-006-9055-4

Schroeder, V. M., & Kelley, M. L. (2009). Associations between family environment, parenting practices, and executive functioning of children with and without ADHD. *Journal of Child and Family Studies*, *18*, 227-235. doi:10.1007/s10826-008-9223-0

Seymour, K. E., Chronis-Tuscano, A., Halldorsdottir, T., Stupica, B., Owens, K., & Sacks, T. (2012). Emotion regulation mediates the relationship between ADHD and depressive symptoms in youth. *Journal of Abnormal Child Psychology*, *40*, 595-606. doi:10.1007/s10802-011-9593-4

Seymour, K. E., Chronis-Tuscano, A., Iwamoto, D. K., Kurdziel, G., & MacPherson, L. (2014). Emotion regulation mediates the association between ADHD and depressive symptoms in a community sample of youth. *Journal of Abnormal Child Psychology*, *42*, 611-621. doi:10.1007/s10802-013-9799-8

Sheehan, K. B., & Pittman, M. (2016). *Amazon's Mechanical Turk for academics: The HIT handbook for social science research*. Irvine, CA: Melvin & Leigh.

Silk, J. S., Steinberg, L., & Morris, A. S. (2003). Adolescents' emotion regulation in daily life: Links to depressive symptoms and problem behavior. *Child Development*, *74*, 1869-1880.

Sjöwall, D., Backman, A., & Thorell, L. B. (2015). Neuropsychological heterogeneity in preschool ADHD: Investigating the interplay between cognitive, affective and motivation-based forms of regulation. *Journal of Abnormal Child Psychology*, *43*, 669-680. doi:10.1007/s10802-014-9942-1

Soenke, M., Hahn, K. S., Tull, M. T., & Gratz, K. L. (2010). Exploring the relationship between childhood abuse and analogue generalized anxiety disorder: The mediating role of emotion dysregulation. *Cognitive Therapy and Research*, *34*, 401-412. doi:10.1007/s10608-009-9264-8

Soto, C. J., John, O. P., Gosling, S. D., & Potter, J. (2008). The developmental psychometrics of big five self-reports: Acquiescence, factor structure, coherence, and differentiation from ages 10 to 20. *Journal of Personality and Social Psychology*, *94*, 718-737. doi:10.1037/0022-3514.94.4.718

Steele, M. M. (2008). Reliability and validity of the Oral Health Scale of the PedsQL: Measuring the relationship between child oral health and health-related quality of life. *Dissertation Abstracts International: Section B. Sciences and Engineering*, *69*, 101.

Steinberg, L. (2005). Cognitive and affective development in adolescence. *Trends in Cognitive Sciences*, *9*(2), 69-74. doi:10.1016/j.tics.2004.12.005

Tull, M. T., Barrett, H. M., McMillan, E. S., & Roemer, L. (2007). A preliminary investigation of the relationship between emotion regulation difficulties and posttraumatic stress symptoms. *Behavior Therapy*, *38*, 303-313. doi:10.1016/j.beth.2006.10.001

U.S. Census Bureau. (2016). *QuickFacts*. Retrieved from https://www.census.gov/quickfacts/fact/table/US/PST045216

Vasilev, C. A., Crowell, S. E., Beauchaine, T. P., Mead, H. K., & Gatzke-Kopp, L. M. (2009). Correspondence between physiological and self-report measures of emotion dysregulation: A longitudinal investigation of youth with and without psychopathology. *Journal of Child Psychology and Psychiatry*, *50*, 1357-1364.

Walden, T. A., Harris, V. S., & Catron, T. F. (2003). How I feel: A self-report measure of emotional arousal and regulation for children. *Psychological Assessment*, *15*, 399-412. doi:10.1037/1040-3590.15.3.399

Wasserstein, R. L. (2016). The ASA's statement on statistical significiance and *P*-values. *The American Statistician*, *1305*. doi:10.1080/00031305.2016.1154108

Weinberg, A., & Klonsky, E. D. (2009). Measurement of emotion dysregulation in adolescents. *Psychological Assessment*, *21*, 616-621.

Wheaton, B., Muthen, B., Alwin, D. F., & Summers, G. F. (1977). Assessing reliability and stability in panel models. *Sociological Methodology*, *8*, 84-136.

Worthington, R. L., & Whittaker, T. A. (2006). Scale development research: Content analysis and recommendations for best practice. *The Counseling Psychologist*, *34*, 806-838. doi:10.1177/0011000006288127

Youngstrom, E. A., Choukas-Bradley, S., Calhoun, C. D., & Jensen-Doss, A. (2015). Clinical guide to the evidence-based assessment approach to diagnosis and treatment. *Cognitive and Behavioral Practice*, *22*, 20-35. doi:10.1016/j.cbpra.2013.12.005

Zimmermann, P., & Iwanski, A. (2014). Emotion regulation from early adolescence to emerging adulthood and middle adulthood: Age differences, gender differences, and emotion-specific developmental variations. *International Journal of Behavioral Development*, *38*, 182-194. doi:10.1177/0165025413515405